

*Konvergenz und Stabilität bei Kollokationsverfahren
vom verallgemeinerten Runge-Kutta Typ*

*Dissertation
zur Erlangung des Grades
Doktor der Naturwissenschaften
am Fachbereich Mathematik
der Johannes Gutenberg Universität in Mainz*

*Peter Fuchs
geboren in Frankfurt/Main*

*Mainz ~~1984~~
1985*



Digitized by the Internet Archive
in 2026

Konvergenz und Stabilität bei Kollokationsverfahren vom verallgemeinerten Runge-Kutta Typ

D I S S E R T A T I O N

zur Erlangung des Grades

D o k t o r d e r N a t u r w i s s e n s c h a f t e n

am Fachbereich Mathematik

der Johannes Gutenberg Universität in Mainz

PETER FUCHS

geboren in Frankfurt/Main

Mainz 1984

Berichterstatter:

Prof. Dr. H. Mülthei, Mainz

Prof. Dr. W. Börsch-Supan, Mainz

Datum der mündlichen Prüfung: 13. Februar 1985

CI STAI



FÜR MEINE SCHWESTER SIGRID
* 1956 + 1981

DEDICATO ALL'ITALIA

ARIÒ

Ariò vuò sapê che d'è
na ventata 'e viento vene n'faccia a te
Ariò è na parola doce
quanno a cante fa parê chiù bella a voce

E' na stella che nun pò carê
ariò ariê
è nun poco 'e bbene pure a te

Ariò è na spia annascusa
che te dice ll'ora bbona p'ogni cosa
Ariò è chi se more 'e sete
passa annanze a na funtana e nun la vede

E' na varca che nun pò affunnà
ariò ariâ
è na storia longa da cuntâ

Facitece verê
ariò ariê
dimme dimme che d' è
Facitece parlâ
ariò ariâ
dimme comme se fa
chi nun o vvo sapê
ariò ariê
io m'o ttengo pe me
e a chi nun vo pazziâ
ariò ariâ
nun c'o vvoglio cantâ

Ariò vuò sapê che fa
ariò te manna 'e nummere a jucâ

E'na ninna nanna nenna nê
ariò ariê
zitto zitto acchiappa pure a me

E si nun ce crire può verê
ariò ariê
quanno te succede pure a te

TERESA DE SIO 1982

Inhalt

1	<u>Einleitung</u>	
1-1	Das untersuchte Verfahren V	- 2-
1-2	Zusammenfassung	- 6-
1-3	Hilfsmittel und Bezeichnungen	- 14-
2	<u>Implizite Kollokationsstellen</u>	
2-1	Darstellungen mittels der Übergangsmatrix der freien Ableitungen	- 18-
2-2	Konvergenz der freien Kollokationsstellen	- 28-
2-3	Darstellungen und Stetigkeitseigenschaften von R	- 34-
3	<u>Ordnung von R und V</u>	
3-1	Charakterisierung der globalen Konvergenz- ordnung von V	- 38-
4	<u>A_{∞}-Stabilität</u>	- 54-
4-1	Charakterisierung der A_{∞} -Stabilität	- 56-
4-2	Zusammenhang zwischen A_{∞} -Stabilität und Konvergenz	- 66-
5	<u>A-Stabilität und Stabilitätsregionen</u>	- 68-
5-1	Existenz A-stabiler Verfahren (Konfluenz)	- 70-
5-2	Beispiele A-stabiler Verfahren	- 82-
5-3	A_{∞} -, A_0 - und A-Stabilität	- 92-
6	<u>Anhang</u>	-100-
	<u>Abbildungen</u>	-108-
	<u>Literaturverzeichnis</u>	-124-

1 Einleitung

1-1 Das untersuchte Verfahren V

Zur numerischen Lösung von Anfangswertproblemen (AWP) bei gewöhnlichen Differentialgleichungen,

$$(1.1) \quad y' = f(x, y), \quad y(0) = y_0, \quad x \in [0, 1],$$

mit hinreichend glattem $f : [0, 1] \times \mathbb{R} \rightarrow \mathbb{R}$ wurde von Mülthei 1979 in [23, 24] und 1980 in [25, 26] ein allgemeines Verfahren beziehungsweise eine Klasse von Verfahren, kurz V genannt, vorgestellt.

Zu gegebenem $n \in \mathbb{N}$ teilt man das zugrundeliegende Intervall $[0, 1]$ in n Teile $J_k := [kh, (k+1)h]$, $k=0(1)n-1$, mit $h := 1/n$ ein. Die Idee des Verfahrens V besteht darin, sukzessive einen Spline $s := s(\cdot, h)$ aus der Stetigkeitsklasse $C^q[0, 1]$ zu erzeugen, der stückweise aus Polynomen $s_k(\cdot, h) \in \Pi_m(J_k)$, $k=0(1)n-1$, zusammengesetzt ist. Dabei sind $m \in \mathbb{N}$, $q \in \mathbb{N}_0$ mit $m > q$ vorgegebene Zahlen.

Die $m+1$ Bedingungen für jedes der Polynome $s_k(\cdot, h)$, $k=0(1)n-1$, ergeben sich aus $q+1$ Stetigkeitsforderungen und $m-q$ weiteren Forderungen, genauer, daß jeder Teilspline $s_k(\cdot, h)$ an $p-1$ Zwischenknoten t_i und im Hauptknoten $t_p := 1$ die Differentialgleichung bis zur d_i -ten Ableitung erfüllt. Dabei ist neben m und q die Anzahl $p \in \mathbb{N}$ der Knoten eine weitere vorgegebene Zahl.

Der polynomiale Höchstgrad genügt demnach mit $d_0 := q$ der Beziehung

$$(1.2) \quad m = \sum_{i=0}^p d_i.$$

In Formeln ist die Erzeugung des Splines s durch (einen Repräsentanten von) V wie folgt implizit definiert:

$$(1.3) \quad \begin{cases} s_k^{(i)}(kh, h) = s_{k-1}^{(j)}(kh, h), \quad j=0(1)q, \\ s_k^{(j)}((k+t_i)h, h) = f^{(j-1)}((k+t_i)h, s_k((k+t_i)h, h)), \\ \quad i=1(1)p, \quad j=1(1)d_i. \end{cases}$$

Dabei bedeutet $f^{(j-1)}$ die $(j-1)$ -te totale Ableitung von f nach dem ersten Argument und es sei

$$(1.4) \quad s_{-1}^{(j)}(0, h) := y^{(j)}(0), \quad j=0(1)q,$$

d.h. man hat q Ableitungen aus der rechten Seite f von (1.1) zu be-

rechnen. Für hinreichend kleines h und glattes f ist s durch (1.3), (1.4) eindeutig bestimmt. In [26] finden sich zwei verschiedene Wege zur praktischen Durchführung des Verfahrens.

Ein jeder Repräsentant von V läßt sich durch einen Vektor

$$(1.5) \quad U := U^{[0]} := (\underbrace{0, \dots, 0}_q, \underbrace{t_1, \dots, t_1}_{d_1}, \dots, \underbrace{1, \dots, 1}_{d_p}) \in \mathbb{R}^m$$

charakterisieren, wobei sinnvollerweise $t_0 := 0$ sei. Im Fall $q > d_p$ soll dieser Vektor stets $U^{[0]}$ heißen, und es werde die Schreibweise $U \in V$ bzw. $U^{[0]} \in V$ verwendet. Damit ist die Verfahrensklasse V mit einer gewissen Menge von Vektoren identifiziert. Zwei Repräsentanten von V , die sich nur in der Nummerierung ihrer Zwischenknoten unterscheiden, sind als gleich anzusehen.

Für die Struktur von V spielt die Größe des Stetigkeitsparameters q in Bezug auf die Vielfachheit d_p im Hauptknoten eine wesentliche Rolle, der Fall " $q \leq d_p$ " ist strukturell einfacher als " $q > d_p$ ". In der vorliegenden Arbeit werden verschiedene, für $q \leq d_p$ bekannte Ergebnisse auf den Fall $q > d_p$ verallgemeinert.

Vor einer näheren Erläuterung soll der Zusammenhang zwischen der Splinekollokationsmethode V und bekannten bzw. neuerdings in der Literatur behandelten Verfahren aufgedeckt werden:

Einerseits sind die Repräsentanten der Teilklasse von V , welche genau die Verfahren vom Typ $q \leq d_p$ enthält und im folgenden $V^{\leq}$ genannt wird, äquivalent zu impliziten Runge-Kutta-Methoden (IRK-Methoden), wenn man von der kontinuierlichen Approximation s aus V nur die diskreten Werte an den Kollokationsknoten betrachtet (Vgl. Mülthei, [23]). Sie von Burage 1978 in [1] diskutierten und besonders für praktische Zwecke gedachten "singly-impliziten RK-Methoden" enthalten eine mit V gemeinsame Teilklasse (Vgl. [1], Bemerkung S.36, (17)), die charakterisiert ist durch:

- (i) $q = d_p = 0$,
- (ii) $d_i = 1, i=1(1)p-1$,
- (iii) $t_i, i=1(1)p-1$, sind Vielfache der Nullstellen gewisser Laguerre-Polynome.

(Man beachte den Sonderfall, daß eine Kollokationsstelle den Wert 1 annimmt ($d_p = 1$).)

Andererseits gibt es einen analogen Zusammenhang zwischen den Repräsentanten der Teilklasse von V , welche die Verfahren vom Typ $q > d_p$ enthält und $V^{>}$ genannt werde, und den von Butcher in [3]

vorgestellten "verallgemeinerten singly-impliziten RK-Methoden". Der Weg, die lokalen Konsistenzordnungen (Konsistenzordnung an den Kollokationsstellen) nicht notwendig so groß zu wählen wie die globale Ordnung, ist auch hier gangbar, was in einer Arbeit von Schneider [40] weiter verfolgt wird. Die von Butcher betrachteten Verfahren bilden dagegen bis auf Sonderfälle, die wieder in $V^{\leq}$ fallen, eine Teilklasse von $V^{>}$, die charakterisiert ist durch:

- (i) $q \geq 1, d_p = 0,$
- (ii) $d_i = 1, i=1(1)p-1,$
- (iii) $t_i, i=1(1)p-1,$ sind Vielfache der Nullstellen gewisser verallgemeinerter Laguerre-Polynome.

1-2 Zusammenfassung

Wie bereits angedeutet, sollen in der vorliegenden Arbeit die Verfahren aus $V^>$ näher untersucht werden. Hier, d.h. bei Kollokationsmethoden mit höheren Ableitungen, welche nicht zu einwertigen IRK-Methoden äquivalent sind, gibt es insbesondere bezüglich A-Stabilität bisher nur vereinzelt zufriedenstellende Resultate, die sich bei Butcher [3] bzw. Janner [44] und Burrage/Chipman [2] finden. Über eine allgemeinere Klasse, nämlich Verfahren, denen bei Anwendung auf die Testgleichung

$$(T_\mu) \quad y' = \mu y, \quad y(0)=1, \quad x \in [0, \infty) \quad ,$$

mit Schrittweite h , $\mu \in \mathbb{C}$, $z := h\mu$, eine charakteristische Relation der Form

$$\phi(\lambda, z) = \sum_{j=0}^M \sum_{i=0}^K a_{ij} z^j \lambda^i = 0$$

mit natürlichen Zahlen M, K und reellen Koeffizienten a_{ij} , $i = 0(1)K$, $j = 0(1)M$, zugeordnet ist¹⁾, existieren verschiedene Untersuchungen von Jeltsch [15-19] und Jeltsch/Nevanlinna [20-21]. Besonders in den beiden zuletzt erwähnten Arbeiten geht es um weitere Aspekte der "linearen Stabilitätstheorie" wie relative Stabilitätsregionen, insbesondere von (expliziten) Taylorpolynomen, Vergleich der Stabilitätsregionen und Ordnungen zweier Verfahren, Konstruktion linearer MSV mit beliebig großer, aber endlicher Stabilitätsregion, Abschätzung von Fehlerkonstanten. A-Stabilität spielt hier eine untergeordnete Rolle. In den erwähnten Arbeiten von Jeltsch ist das Resultat bzgl. A-Stabilität, welches auf eine größere Klasse als die der linearen MSV anwendbar ist, negativ ([15], Theorem 2). Konkrete Resultate bzgl. dieser wichtigen Eigenschaft bei Kollokationsmethoden finden sich in [2], [3], [44].

In der vorliegenden Arbeit soll daher insbesondere angestrebt werden, Parameter t_i, d_i , $i = 1(1)p-1$, explizit zu bestimmen derart, daß für das zugeordnete $U^{(0)} \in V^>$ A-Stabilität vorliegt, mit anderen Worten einfach überprüfbare Kriterien für diese Eigenschaft zu erstellen.

Die komplizierter Struktur des Verfahrens für $q > d_p$ legt die Frage nahe, warum man diesen Fall überhaupt untersucht, wo es doch

¹⁾ Siehe Jeltsch/Nevanlinna [21], S.248. Für die Verfahren aus der Klasse $V^>$ wird diese Zuordnung hier in Theorem 2.2 gezeigt.

Für praktische Zwecke schon zufriedenstellende Resultate für die Klasse $V^{\leq}$ gibt. Eine Antwort darauf ist, daß man - vgl. (1.2) - die Ordnung des Verfahrens, im allgemeinen gleich m , entweder über den Parameter q oder über die Größen d_i , $i=1(1)p$, erhöhen kann. Je Erhöhung der Ordnung um 1 muß man im ersten Fall in jedem Schritt der Verfahrens den zuletzt berechneten Teilspline $s_{k-1}(.,h)$, also ein Polynom, einmal mehr ableiten und auswerten, im zweiten Fall dagegen in jedem Schritt die rechte Seite f aus (1.1) einmal mehr (ableiten und) auswerten, was im allgemeinen einen höheren Aufwand bedeutet.

Vor der genaueren Erläuterung des Inhaltes der vorliegenden Arbeit zunächst einige Bemerkungen zu bekannten Resultaten im Fall $q \leq d_p$: Eine entscheidende Rolle für Aussagen über Konvergenz und Stabilität spielt hier eine rationale Funktion R , die sogenannte Wachstumsfunktion, welche man erhält, wenn man $U \in V^{\leq}$ auf die Testgleichung (T_μ) , $\mu \in \mathbb{C}$, mit fester Schrittweite $h > 0$ anwendet und als Argument $z := h\mu$ setzt. Dann gilt nämlich mit dieser Funktion R und

$$(1.6) \quad \begin{aligned} s_k &:= s_k(z) := s(kh) & : \\ s_k(z) &= (R(z))^k, \quad k \in \mathbb{N}_0. \end{aligned}$$

Wanner und Nørsett haben 1979 in [32] die Zuordnung

$$\tilde{Q} : \begin{cases} V^{\leq} \longrightarrow \mathbb{C}(z) \\ U \longmapsto \tilde{Q}(.,U) \end{cases}$$

explizit angegeben, für die

$$(1.7) \quad R(z) = \tilde{Q}(z,U), \quad z \in \mathbb{C},$$

gilt. Die Bezeichnung $\tilde{Q}$ wurde gewählt, weil zwischen dieser Funktion und dem Spektralradius einer gewissen matrixwertigen Funktion Ψ ein Zusammenhang besteht, was später einsichtig wird (Satz 2.10). Der Buchstabe Q sei dem Spektralradius selbst vorbehalten. Man kann nun beweisen, daß die Funktion R die Konvergenzordnung ν des Verfahrens $U \in V^{\leq}$ charakterisiert. Insbesondere gilt (Vgl. z.B. Burrage [1], Korollar 2.1, Nørsett/Wanner [32], Theorem 4, Mülthei [28], (4.5)):

$$(1.8) \quad \nu = m+1 \iff \int_0^1 \tau^q \prod_{i=1}^p (\tau - t_i)^{d_i} d\tau = 0.$$

Desweiteren läßt sich die A_∞ -Stabilität mittels der Größe

$$(1.9) \quad R(\infty) = \begin{cases} 0 & , q < d_p, \\ (-1)^{m-d_p} \prod_{i=1}^{p-1} \left(\frac{1-t_i}{t_i} \right)^{d_i} & , q = d_p, \end{cases}$$

wie folgt charakterisieren: U ist genau dann A_∞ -stabil, wenn $|R(\infty)| < 1$ gilt. Schließlich sind wegen

$$(1.10) \quad R_A = \{ z \in \mathbb{C} \mid |R(z)| < 1 \}$$

unter Anwendung des Maximumsprinzips auf (R, G) mit $G := \mathbb{C}^- \subset \mathbb{K} := \mathbb{C} \cup \{\infty\}$ und unter Ausnutzung der bekannten Struktur von R Untersuchungen möglich, die die explizite Angabe von Parametern erlauben, für die A-Stabilität vorliegt (Vgl. z.B. Mülthei, 1982, [29]). Hierbei bezeichne R_A die Region absoluter Stabilität von U. Im übrigen existieren zahlreiche auch neuere Arbeiten über die rationale Approximation der Exponentialfunktion, beispielsweise die Publikationen von Hairer [12], Nørsett [30-31], Nørsett/Wolfbrand [33], Saff/Varga [37-39], Wanner [42-43].

Nachfolgend soll der Inhalt der vorliegenden Arbeit etwas näher ausgeführt werden. Bei Anwendung eines $U^{[0]} \in V^>$ auf (T_μ) in der oben beschriebenen Weise ergibt sich mit

$$(1.11) \quad \begin{aligned} s_k &:= s_k(z) := s(kh) : \\ s_k(z) &= R_{k-1}(z) \cdot \dots \cdot R_0(z)^{-1}, \quad k \in \mathbb{N}_0, \end{aligned}$$

wobei die R_j , $j \in \mathbb{N}_0$, verschiedene rationale Funktionen sind. Grundlegende Idee der vorliegenden Arbeit ist es, die Funktion

$$(1.12) \quad R := \lim_{k \rightarrow \infty} R_k$$

punktweise zu definieren und mit deren Hilfe analoge Eigenschaften wie im Fall $q \leq d_p$ zu beweisen. Es stellt sich heraus, daß $R(z)$ für alle z außerhalb einer gewissen Ausnahmemenge -im allgemeinen eine

¹⁾ Durch das Verfahren ist zunächst $(s_k(z))_{k \in \mathbb{N}_0}$, $z \in \mathbb{C}$ -Pole von R_0 , gegeben und man definiert R_j vermöge $R_j(z) := s_{j+1}(z)/s_j(z)$, $j \in \mathbb{N}_0$ (Siehe Definition 2.3).

Menge vom Maße Null- betragsgröÙter Eigenwert einer Matrix $\Psi(z)$ ist, welche man aus der matriziellen Beschreibung des Verfahrens (Vgl. Mülthei [27], Satz 1) erhalten kann.

Im folgenden vorbereitenden Kapitel wird gezeigt, daÙ das in (1.12) definierte Grenzelement formal analog zu (1.7) dargestellt werden kann, sofern man nur im Vektor $U^{[0]}$ die ersten $q-d_p$ Nullen durch gewisse Funktionen $u_j : \mathbb{C} - F \rightarrow \mathbb{C}$, $j=1(1)q-d_p$, ersetzt. Dann gilt für die Größe R aus (1.12) :

$$(1.13) \quad \left\{ \begin{array}{l} R(z) = \tilde{g}(z, U(z)) \quad , \text{ mit} \\ U(z) = (u_1(z), \dots, u_{q-d_p}(z), \underbrace{0, \dots, 0}_{d_p}, \underbrace{t_1, \dots, t_1}_{d_1}, \dots, \underbrace{1, \dots, 1}_{d_p}) \\ \qquad \qquad \qquad \in \mathbb{C}^m, \quad z \in \mathbb{C} - F \quad . \end{array} \right.$$

Hierbei ist F eine Ausnahmемenge, die sich im allgemeinen aus einer endlichen Punktmenge und eindimensionalen komplexen Mannigfaltigkeiten (Linien) zusammensetzt. Die Funktionen u_j , $j=1(1)q-d_p$, lassen sich als implizite Kollokationsstellen deuten. Im Fall $q=d_p+1$ kann man u_1 explizit, im allgemeinen aber nur implizit angeben, was jedoch für die angestrebten Aussagen genügt. Mit (1.13) ist eine Darstellung von R gegeben, die mit der natürlichen matriziellen Beschreibung des Verfahrens auf den ersten Blick nicht mehr viel zu tun hat, später jedoch wesentlich eingeht. Um diese Struktur zu beweisen, sind die freien Ableitungen

$$s_k^{(j)}(kh, h), \quad j=d_p+1(1)q,$$

die den Übergangsbedingungen aus (1.3), erste Zeile, genügen, beziehungsweise die freien Nullstellen von

$$s_k'(\cdot, h) - \mu s_k(\cdot, h)$$

zu untersuchen. Dabei stellt sich heraus, daÙ die z - und k -abhängigen Koeffizientenvektoren der Polynome mit Hauptkoeffizient 1, die den freien Nullstellen zugeordnet werden, beim Grenzübergang $k \rightarrow \infty$ gegen einen Eigenvektor zum Eigenwert $R(z)$ der Matrix $\Psi(z)$ konvergieren.

Das dritte Kapitel der Arbeit soll von Charakterisierung der Konvergenzordnung handeln. Die Charakterisierungsaussage stellt sich im wesentlichen als eine Folgerung aus der Darstellung (1.13) her-

aus. Dabei lassen sich sogar die Fälle Divergenz bzw. Konvergenzordnung $\nu = m-1$ ¹⁾ erfassen. Die Integralbedingung für die maximale Konvergenzordnung $\nu = m+1$, nämlich

$$(1.14) \quad \int_0^1 \prod_{j=1}^{q-d_p} (\tau - u_j(0)) \tau^{d_p} \prod_{i=1}^p (\tau - t_i)^{d_i} d\tau = 0,$$

die man formelmäßig auch aus der Rangbedingung (4.4) in Mülthei [28] berechnen kann, erhält hier durch die mögliche Interpretation der Zahlen $u_j(0)$, $j=1(1)q-d_p$, "mehr Farbe". Darüber hinaus erlaubt die Darstellung (1.13) eine anschauliche Erklärung insbesondere für den von Mülthei in [24], Satz 5 und Nachtrag, bzw. [25], Satz 2, beschriebenen seltsamen Effekt, nämlich, daß im Fall $p=1$, $q=d_p+1$, für die Konvergenzordnung

$$\nu = \begin{cases} m-1 & , d_p \text{ gerade,} \\ m+1 & , d_p \text{ ungerade,} \end{cases}$$

gilt. Die Frage, ob auch für Verfahren aus $V^>$ eine höhere Ordnung als $m+1$ (in den Hauptknoten) möglich ist, besitzt bislang erst für die Anwendung auf (T_μ) eine positive und konstruktive Antwort, weswegen hierauf nicht näher eingegangen werden soll.

Im vierten Kapitel wird A_∞ -Stabilität untersucht. Bekanntlich sind A_∞ -stabile Verfahren in vielen Fällen "fast A-stabil". Hier läßt sich eine Gleichung der Form

$$f_\infty(\lambda) = \lambda^{q-d_p+1} + a_{q-d_p}(\infty) \lambda^{q-d_p} + \dots + a_0(\infty) = 0$$

mit gewissen reellen, explizit in Abhängigkeit der Verfahrensparameter bekannten Zahlen $a_j(\infty)$, $j=0(1)q-d_p$, angeben, die die folgende Bedeutung hat: A_∞ -Stabilität liegt im allgemeinen genau dann vor, wenn sich sämtliche Nullstellen obiger Gleichung im offenen Einheitskreis befinden (Transformation, Routh-Hurwitz-Kriterium!). Mit

$$\beta := (-1)^{m-d_p} \prod_{i=1}^{p-1} \left(\frac{t_i}{1-t_i} \right)^{d_i}, \quad \Gamma := 2d_p + \sum_{i=1}^{p-1} \frac{d_i}{t_i(1-t_i)}$$

lautet obige Gleichung im Fall $q = d_p+1$: $\lambda^2 + (2+\Gamma)/\beta \cdot \lambda + 1/\beta^2 = 0$.

1) Hier nur für $q=d_p+1$ untersucht. Def. der Konvergenz s. Kapitel 3.

Auch hier läßt sich analog zu (1.9) die Größe

$$(1.15) \quad R(\infty) = \frac{1}{\beta} \cdot \frac{1 + \sqrt{1+4/\Gamma}}{1 - \sqrt{1+4/\Gamma}}$$

angeben, so daß man A_∞ -Stabilität -für $|R(\infty)| \neq 1$ - entsprechend dem Fall $q \leq d_p$ charakterisieren kann: $U^{[0]}$ ist genau dann A_∞ -stabil, wenn $|R(\infty)| < 1$ gilt. In Spezialfällen kann man die Parameterintervalle, für die A_∞ -Stabilität vorliegt, explizit bestimmen. Die Untersuchungen haben einen weiteren, unerwarteten Zusammenhang ergeben. Es gilt nämlich (fast stets):

$$U^{[0]} \text{ ist } A_\infty\text{-stabil} \Leftrightarrow (0, U^{[0]}) \text{ ist konvergent.}$$

Insbesondere lassen sich mit der im vierten Kapitel genau beschriebenen Methode zur Bestimmung der Parameterintervalle mit A_∞ -Stabilität analog Konvergenzbereiche im Parameterraum ermitteln.

Im Kapitel 5 schließlich werden die für die praktischen Anwendungen wichtigen Stabilitätsregionen und deren Eigenschaften, insbesondere A_0 - und A -Stabilität untersucht. Es gilt (1.10) auch für das in (1.12) definierte R . Zunächst könnte man vermuten, daß es in $V^>$ keine A -stabilen Verfahren gibt, da die Funktion R_0 aus (1.11) in $z=\infty$ einen Pol hat. Dies erweist sich als Trugschluß, wenn auch zu bemerken ist, daß absolute Stabilität in einem Punkte $z \in \mathbb{C}$ im Fall $q > d_p$ eine (etwas) andere Bedeutung als für $q \leq d_p$ hat, was später erläutert wird. Mittels der Idee der Konfluenz (Man bewegt sich bzgl. des Parametersatzes $U^{[0]} \in V^>$ "in der Nähe" eines Verfahrens aus $V^{\leq}$) gelingt es, die Existenz A -stabiler Verfahren in $V^>$ sehr allgemein zu beweisen. Dieselbe Idee liefert dann auch die Möglichkeit, mittels Abschätzungen Parameterintervalle konstruktiv zu bestimmen, für die A -Stabilität vorliegt. Dabei gehen zwei Eigenschaften der Funktion R wesentlich ein, einmal die Struktur (1.13), zum anderen die Tatsache, daß $R(z)$, wie erwähnt, für fast alle $z \in \mathbb{C}$ betragsgrößerer Eigenwert einer Matrix $\Psi(z)$ ist. Es folgen Beispiele und ein Vergleich mit den Ergebnissen über A -Stabilität von Butcher [3] bzw. Wanner [44] und Burrage/Chipman [2]. Bei der hier verfolgten Vorgehensweise stellt sich heraus, daß man auch für $V^>$ in großen Parameterintervallen A -Stabilität erreichen kann, wenn man nur die Kollokationsstellen t_i in geeigneter Weise "um die 1" anordnet. Man kann sich mit den Parametern sogar soweit von 1 entfernen, daß

man eigentlich nicht mehr von Konfluenz reden möchte. Ein Vergleich der numerisch bestimmten "genauen" Parameterintervalle, für die A-Stabilität vorliegt, mit den Bereichen mit A_0 - bzw. A_∞ -Stabilität bei "obigen" Beispielen schließt die Arbeit und führt zu der unbewiesenen Beobachtung, daß man in einigen Fällen die Grenzen der "A-Stabilitätsintervalle" durch Nullstellen angegebbarer Polynome charakterisieren kann.

■ 1-2

Mein besonderer Dank gebührt dem Betreuer dieser Arbeit, Herrn Prof.Dr. H. Mülthei, nicht alleine für die Anregung zu dem interessanten Thema, sondern auch für sein starkes Engagement für formale Klarheit, was wesentlich dafür war, daß die vorliegende Arbeit in der jetzigen Form entstehen konnte. Hierzu hat nicht zuletzt auch die Arbeitsgruppe Numerische Mathematik durch zahlreiche Diskussionen beigetragen.

Ebenso anerkenne ich die freundliche Unterstützung von Herrn Prof.Dr. E. Gottschling, der in einigen Gesprächen bei der Klärung funktionentheoretischer Aspekte dieser Arbeit geholfen hat.

1-3 Hilfsmittel und Bezeichnungen

1.1 Definition

a) Seien $M \in \mathbb{N}$, $w = (w_1, \dots, w_M) \in \mathbb{C}^M$. Man definiere (wie üblich) die symmetrischen Grundfunktionen $\sigma_1 : \mathbb{C}^M \rightarrow \mathbb{C}$ vermöge

$$(1.16) \quad \sigma_1(w) := \begin{cases} \sum_{1 \leq i_1 < i_2 < \dots < i_l \leq M} w_{i_1} \cdot \dots \cdot w_{i_l}, & l=1(1)M, \\ 1, & l=0, \\ 0, & \text{sonst.} \end{cases}$$

b) Zu $M \in \mathbb{N}$, $w = (w_1, \dots, w_M) \in \mathbb{C}^M$ sei $w^* := (1, \dots, 1) - w$.

c) Seien $\mathbb{K} := \mathbb{C} \cup \{\infty\}$, $f : \mathbb{K} \rightarrow \mathbb{K}$, $N_f := \{z \in \mathbb{K} \mid f(z) = 0\}$,

$$P_f := \{z \in \mathbb{K} \mid f(z) = \infty\}.$$

d) Zu $U^{(0)} \in V^>$ (Vergleiche (1.5)) seien

$$\begin{aligned} T &:= (\underbrace{t_1, \dots, t_1}_{d_1}, \dots, \underbrace{t_{p-1}, \dots, t_{p-1}}_{d_{p-1}}, t_{p-1}) \in \mathbb{R}^{m-q-d_p}, \\ U_2^{(1)} &:= (\underbrace{0, \dots, 0}_{d_p}, \underbrace{T, 1, \dots, 1}_{d_p}) \in \mathbb{R}^{m-(q-d_p)}. \end{aligned}$$

e) Seien $M \in \mathbb{N}$, $A \in L(\mathbb{C}^M)$. $S(A)$ bezeichne das Spektrum der Matrix A , $\rho(A)$ deren Spektralradius.

1.2 Theorem (Nørsett/Wanner)

Seien $M \in \mathbb{N}$, $w = (w_1, \dots, w_M) \in \mathbb{C}^M$,

$$(1.17) \quad \left\{ \begin{aligned} n(t, w) &:= n(t, w, M) := \frac{1}{M!} \prod_{i=1}^M (t - w_i), & t \in \mathbb{R}, \\ \tilde{p}(z, w) &:= \tilde{p}(z, w, M) := \sum_{l=0}^M n^{(M-1)}(1, w) z^l, \\ \tilde{q}(z, w) &:= \tilde{q}(z, w, M) := \sum_{l=0}^M n^{(M-1)}(0, w) z^l, \\ \tilde{g}(z, w) &:= \tilde{g}(z, w, M) := \tilde{p}(z, w) / \tilde{q}(z, w), & z \in \mathbb{K}. \end{aligned} \right.$$

Dann gelten die folgenden Aussagen :

¹⁾ Definition von U_1 erfolgt später (Vergleiche (1.13), (2.17)).

a)

$$e^z - \tilde{g}(z, w) = \frac{z^{M+1}}{\tilde{q}(z, w)} \int_0^1 e^{z(1-\tau)} n(\tau, w) d\tau, \quad z \in \mathbb{C}.$$

b) Für beliebiges $U^{[0]} \in V$ gilt bei Anwendung auf (T_μ) mit fester Schrittweite $h > 0$, $\mu \in \mathbb{C}$, $z := h\mu$, nach dem ersten Schritt:

$$\tilde{g}(z, U^{[0]}) = s_1 = \begin{cases} R(z) & , U^{[0]} \in V^{\leq}, \\ R_0(z) & , U^{[0]} \in V^>, \quad z \in \mathbb{C} - N_{\tilde{q}}(\cdot, U^{[0]}) \end{cases}$$

Beweis

Nörsett/Wanner [32], Theorem 4.

1.3 Lemma

Seien $M \in \mathbb{N}$, $w = (w_1, \dots, w_M) \in \mathbb{C}^M$, $J \subset \{1, \dots, M\}$, $w' := (w_j), j \notin J^{(1)} \in \mathbb{C}^{M-\#(J)}$. Dann folgt:

$$\lim_{\substack{w_j \rightarrow \infty \\ j \in J}} \tilde{g}(z, w, M) = \tilde{g}(z, w', M-\#(J)), \quad z \in \mathbb{K}.$$

Bemerkung

Der obige Limes erlaubt es, in w formal ∞ als Kollokationspunkt zuzulassen. Man könnte etwa

$$\tilde{g}(z, (w', \underbrace{\infty, \dots, \infty}_M), M) := \tilde{g}(z, w', M-\#(J))$$

setzen. Die anschauliche Bedeutung dieser Tatsache liegt darin, daß "weit außerhalb" liegende Kollokationsstellen für die Wachstumsfunktion und damit für die (analytischen) Stabilitätseigenschaften des Verfahrens keine Rolle spielen. Diese Eigenschaft liefert später die in 1-2 angesprochene anschauliche Erklärung für den seltsamen Effekt von Konvergenzordnung $m-1$ bzw. $m+1$ im Fall $p=1, q=d_p+1$.

Beweis

Mit (1.17) folgen die beiden Beziehungen:

¹⁾ Die Schreibweise $(w_j), j \notin J$ bezeichne denjenigen Vektor, den man erhält, wenn man aus w die Komponenten mit Indizes in J herausstreicht.

$$\begin{aligned} \text{grad } \tilde{p}(.,w) &= M - \#(\{j \in \{1, \dots, M\} \mid w_j = 1\}), \\ \text{grad } \tilde{q}(.,w) &= M - \#(\{j \in \{1, \dots, M\} \mid w_j = 0\}). \end{aligned}$$

Die Grade ändern sich nicht beim angegebenen Grenzübergang, was die Aussage für $z=\infty$ insbesondere bei gleicher Anzahl von Nullen und Einsen im Vektor w liefert. Sei $z \in \mathbb{C} - \{0\}$. Lemma 6.1.c und f aus dem Anhang liefern eine Darstellung für $\tilde{q} = \tilde{p}/\tilde{q}$ in Abhängigkeit von symmetrischen Funktionen von w . Zähler und Nenner werden mit

$$(-z)^{\#(J)} \prod_{j \in J} w_j$$

gekürzt. Die Behauptung folgt unter Verwendung von Lemma 6.1.a mit $w_1 := (w_j), j \in J, w_2 := w'$, nach einiger Rechnung.

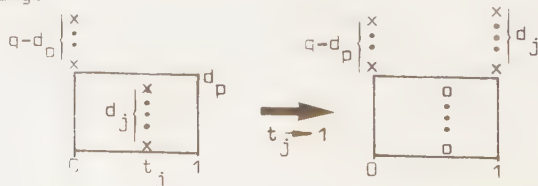
Für Zitate innerhalb der vorliegenden Arbeit werden ab sofort die Abkürzungen B (Bemerkung(en)), D (Definition), K (Korollar), L (Lemma), S (Satz) sowie T (Theorem) verwendet.

1.4 Bemerkungen

a) Folgende Skizze möge dazu dienen, den Parametersatz (zu einem Verfahren) $U^{[0]} \in V^>$ etwas anschaulicher darzustellen.



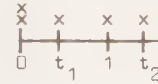
$(R^{[0]} := \tilde{q}(z, U^{[0]}, m))$ Hierbei bedeute das Rechteck irgendeinen festen Parametersatz, der konkret nicht weiter interessiert, wobei $t_1 > 1$ und $d_1 > d_p$ durchaus zulässig sein soll. Die Vielfachheiten in 0 bzw. 1 sind q bzw. d_p . Interessiert ein spezielles t_j besonders, so wird dies gesondert einschließlich seiner Vielfachheit eingezeichnet. Bei Grenzübergängen, die später eine wichtige Rolle spielen, diene die folgende Skizze der Erläuterung:



Die "Löcher" im Rechteck auf der rechten Seite sollen andeuten, daß (gegenüber vorher) an dieser Stelle "etwas fehlt".

- b) Bei konkreten Beispielen werden sämtliche Kollokationsknoten angedeutet und deren Vielfachheiten angegeben. Beispiel:

$$p=3, q=2, d_3=1; d_1=d_2=1, t_1 \in (0,1), t_2 > 1.$$



1.5 Bezeichnung

Im folgenden spielen Vektoren aus $\mathbb{C}^{q-d_p+1}$ und Matrizen aus $L(\mathbb{C}^{q-d_p+1})$ eine wichtige Rolle. Dabei ist es natürlich, die Komponenten von 0 bis $q-d_p$ beziehungsweise von 0,0 bis $q-d_p, q-d_p$ durchnummerieren. Ferner seien zu $x = (x_0, \dots, x_{q-d_p})^T$

$$x_r := (x_1, \dots, x_{q-d_p})^T$$

und zu einer Matrix $A = (a_{ij}), i, j=0(1)q-d_p,$

$$A_r := (a_{ij}), i, j=1(1)q-d_p,$$

definiert. Der Buchstabe r soll an reduziert erinnern und wird im Zusammenhang mit solchen Größen nicht als Index verwendet. Indizes, die Komponenten von Vektoren oder Matrizen bezeichnen, stehen grundsätzlich rechts unten am Namen der zugehörigen Größe.



2 Implizite Kollokationsstellen

2-1 Darstellungen mittels der Übergangs-
matrix der freien Ableitungen

Im folgenden sei, wenn nicht anders vermerkt, ein beliebiges, aber festes $U^{[0]} \in V^>$ vorgegeben. Für das gesamte Kapitel 2 gelte, daß $U^{[0]}$ stets auf (T_μ) , $\mu \in \mathbb{C}$ beliebig, mit konstanter Schrittweite $h > 0$ angewendet wird, und es sei wie "oben" $z := h\mu$ gesetzt. In diesem Kapitel wird, in Verallgemeinerung von Aussagen für den Fall $q \leq d_p$, eine Funktion R untersucht, deren (strukturelle) Eigenschaften später, nämlich bei der Charakterisierung der Konvergenzordnung und Untersuchung der Stabilitätsregionen, wesentlich eingehen.

2.1 Definition

Sei $U^{[0]} \in V$. Mit der Menge

$$R_A := R_A(U^{[0]}) := \left\{ z := h\mu \in \mathbb{C} \mid U^{[0]} \text{ angewandt auf } (T_\mu) \text{ mit konstanter Schrittweite } h > 0 \text{ liefert eine Folge } (s_k(z))_{k \in \mathbb{N}_0} \text{ mit } \lim_{k \rightarrow \infty} s_k(z) = 0. \right\}$$

werde die Region absoluter Stabilität (von $U^{[0]}$) bezeichnet.

Um R_A zu untersuchen, benötigt man eine explizite Darstellung der Größen $s_k(z)$, $k \in \mathbb{N}_0$, in Abhängigkeit der Verfahrensparameter. Diese wird in der folgenden Aussage gegeben, wobei die Übergangsmatrix der Vektoren, die aus dem Funktionswert s_k und den freien Ableitungen gebildet werden, die wesentliche Rolle spielt.

2.2 Theorem

Sei $U^{[0]} \in V^>$. Dann gilt in Verallgemeinerung von (1.6):

$$(2.1) \quad s_k(z) = \frac{q(z)^T \Psi(z)^k e_0}{q_0(z)}, \quad z \in \mathbb{C} - N_{q_0}, \quad k \in \mathbb{N}_0.$$

Hierbei sind $q = (q_0, \dots, q_{q-d_p})^T : \mathbb{C} \rightarrow \mathbb{C}^{q-d_p+1}$ und die matrixwertige Funktion $\Psi : \mathbb{C} - N_{q_0} \rightarrow L(\mathbb{C}^{q-d_p+1})$ wie folgt gegeben:

$$(2.2) \quad \left\{ \begin{array}{l} \psi(z) = \begin{pmatrix} q(z)^T \\ 0, H \end{pmatrix}^{-1} \begin{pmatrix} p(z)^T \\ h^*, H^* \end{pmatrix}, z \in \mathbb{C} \setminus N_{q_0}, \text{ mit} \\ p(z) = (p_0(z), \dots, p_{q-d_p}(z))^T, \\ |p_\alpha(z)| := (-z)^\alpha \frac{(m-\alpha)!}{m!} \tilde{p}(z, U^{[\alpha]}, m-\alpha), \\ q_\alpha(z) := (-z)^\alpha \frac{(m-\alpha)!}{m!} \tilde{q}(z, U^{[\alpha]}, m-\alpha), z \in \mathbb{C}, \\ U^{[\alpha]} := (\underbrace{0, \dots, 0}_{q-\alpha}, \underbrace{1, \dots, 1}_{d_p}), \alpha = 0(1)q-d_p, \end{array} \right.$$

sowie $H, H^* \in GL(\mathbb{R}^{q-d_p})$, $h^* \in \mathbb{R}^{q-d_p}$, definiert durch

$$(2.3) \quad \left\{ \begin{array}{l} (0, H) := \begin{pmatrix} v^{(d_p)}(0)^T \\ \vdots \\ v^{(q-1)}(0)^T \end{pmatrix}, (h^*, H^*) := \begin{pmatrix} v^{(d_p)}(1)^T \\ \vdots \\ v^{(q-1)}(1)^T \end{pmatrix}, \\ v(t) := t^{d_p} \prod_{i=1}^p (t-t_i)^{d_i} t^{q-d_p}, \dots, (-1)^{q-d_p} t^0)^T, \\ t \in \mathbb{R}. \end{array} \right.$$

Bemerkungen

a) Gemäß (1.17) und T 1.2.b ist der z-abhängige Teil in der Matrix $\psi(z)$ aus denjenigen Funktionen aufgebaut, die sich bei Integration der Testgleichung von 0 bis h mit den "artverwandten" Verfahren

$$U^{[\alpha]} = \begin{matrix} & & x \\ & & \vdots \\ q-d_p-\alpha & \left\{ \begin{matrix} \vdots \\ x \end{matrix} \right. & & \\ & \boxed{} & d_p & \\ & & & \end{matrix} \quad , \quad \alpha = 0(1)q-d_p,$$

ergeben. Eine Darstellung von $\psi(z)$ ausschließlich mit symmetrischen Grundfunktionen und Fakultäten sowie einige Eigenschaften dieser Matrix, die später benötigt werden, finden sich im Anhang (S 6.3).

b) Die von Wanner [44] für den Spezialfall $q=1$, $d_p=0$, $d_i=1$, $i=1(1)p-1$, angegebene Matrix $R(z)$ ist ähnlich zu $\psi(z)$, wie man nachrechnen kann. Dasselbe gilt auch bezüglich der Matrix

$R + zB(I - zA)^{-1}Q$, die Butcher in [3] für verallgemeinerte (singly-implizite) RK-Methoden angibt, was aufgrund der Einbettung dieses Verfahren in die Klasse $V^>$ klar ist (Anwendung von (2.1), (2.2) in [3] auf (T_μ) , $z := h\mu$, Eliminierung der Größen $y_1, \dots, y_s$; siehe auch Korollar im Anschluß an den Beweis des hiesigen Theorems).

- c) Wegen (2.1) kann man sich die Werte $(s_k(z))_{k \in \mathbb{N}_0}$ auch durch eine lineare Differenzengleichung erzeugt denken. Die Eigenwerte $\lambda := \lambda(z)$ dieser Gleichung, mit anderen Worten die Eigenwerte von $\Psi(z)$, genügen -vgl. L 6.2- der folgenden Relation:

$$(2.4) \quad \begin{aligned} f_{\Psi(z)}(\lambda) = & \det(\lambda I_r - H^{-1}H^*) \left[\lambda - \frac{1}{q_0(z)}(p_0(z) - q_r^T(z)H^{-1}h^*) \right. \\ & \left. + \frac{1}{q_0(z)}(q_r^T(z)H^{-1}h^* - p_r^T(z))(\lambda I_r - H^{-1}H^*)^{-1}H^{-1}h^* \right] = 0 \end{aligned}$$

Die Untersuchung der Nullstellen dieser (algebraischen) Gleichung im Fall der Konfluenz (Siehe Einleitung) wird in Kapitel 5 eine wichtige Rolle spielen.

- d) Aufgrund der trivialen Folgerung

$$R_A \supset S := \{ z \in \mathbb{C} \mid \rho(\Psi(z)) < 1 \}$$

besteht bereits eine Möglichkeit, Regionen absoluter Stabilität zu untersuchen, was jedoch ohne weitere Kenntnis über $\rho(\Psi(z))$ problematisch sein dürfte. Zumindest für $q = d_p + 1$ lassen sich $\#(R_A - S) < \infty$ und die praktische Unbrauchbarkeit der Punkte aus $R_A - S$ zeigen. Für alle $z \in S$ jedoch bewirkt eine geringfügige Störung der Startwerte $y^{(j)}(0)$, $j = d_p + 1(1)q$, der freien Ableitungen nichts wesentliches, die Eigenschaft $\lim_{k \rightarrow \infty} s_k(z) = 0$ bleibt erhalten.

Beweis

Der Beweis geht auf die ursprüngliche Formulierung (1.3), (1.4) des Verfahrens zurück, wobei die Untersuchung der freien Ableitungen $s_k^{(j)}(kh, h)$, $j = d_p + 1(1)q$, wesentlich eingeht. Hierbei werden die Ideen aus dem Beweis von Nørsett/Wanner [32], Theorem 4, verallgemeinert. Im Verlauf des hiesigen Beweises zeigt sich bei Auflösung eines linearen Gleichungssystems, daß die Vektorenfolge

$$((s_k(z), h^{d_p+1} s^{(d_p+1)}(kh, h), \dots, h^{q_p} s^{(q)}(kh, h))^T)_{k \in \mathbb{N}_0}$$

und damit -wie man sich Überlegen muß- auch die Folge

$(s_k(t, h), t \in [kh, (k+1)h])_{k \in \mathbb{N}_0}$ für jedes $z \in \mathbb{C} - N_{q_0}$ eindeutig bestimmt ist. Dies werde der Einfachheit halber vorweggenommen.

Es seien

$$(2.5) \quad \begin{cases} \omega_k(t) := s_k(th, h), & t \in [k, k+1], \\ \phi_k(t, z) := \omega_k'(t+k) - z\omega_k(t+k), & t \in [0, 1], z \in \mathbb{C} - N_{q_0}, k \in \mathbb{N}_0. \end{cases}$$

Aus dem Beweis zu Theorem 4 in [32] ergibt sich

$$\phi_0(t, z) = \frac{-z^{m+1}}{m!q_0(z)} v(t)^T e_0, \quad t \in [0, 1], z \in \mathbb{C} - N_{q_0}.$$

Da $\phi_k(\cdot, z)$ für jedes $k \in \mathbb{N}_0$ ein Polynom vom Höchstgrad m ist, kann man unter Berücksichtigung der Kollokationsbedingungen (1.3) den Ansatz

$$(*) \quad \phi_k(t, z) = \frac{-z^{m+1}}{m!q_0(z)} v(t)^T \chi^{[k]}(z), \quad t \in [0, 1], z \in \mathbb{C} - N_{q_0}, k \in \mathbb{N}_0,$$

mit v aus (2.3) erstellen, wobei $\chi_k := \chi^{[k]}(z)$ gesucht ist. Man überlegt sich leicht die folgende Relation:

$$(**) \quad \omega_k(t+k) = -\sum_{l=0}^m z^{-(l+1)} \phi_k^{(l)}(t, z), \quad t \in [0, 1], z \in \mathbb{C} - N_{q_0}, k \in \mathbb{N}_0.$$

Die noch nicht in den Ansatz (*) eingebrachten Bedingungen aus (1.3), nämlich $\omega_k(k) = \omega_{k-1}(k)$ sowie $\phi_k^{(j-1)}(0, z) = \phi_{k-1}^{(j-1)}(1, z)$, $j = d_p + 1(1/q)$, liefern mittels (**), (*) das lineare Gleichungssystem

$$\frac{-z^{m+1}}{m!q_0(z)} \begin{pmatrix} -\sum_{l=0}^m z^{-(l+1)} v^{(l)}(0)^T \\ v^{(d_p)}(0)^T \\ \vdots \\ v^{(q-1)}(0)^T \end{pmatrix} \chi_k = \frac{-z^{m+1}}{m!q_0(z)} \begin{pmatrix} -\sum_{l=0}^m z^{-(l+1)} v^{(l)}(1)^T \\ v^{(d_p)}(1)^T \\ \vdots \\ v^{(q-1)}(1)^T \end{pmatrix} \chi_{k-1}.$$

Die Formeln (2.3) sowie (1.17) besagen für die Komponenten von $v(t)$:

$$v_\alpha(t) = (-1)^\alpha t^{q-\alpha} \prod_{i=1}^p (t-t_i)^{d_i}$$

$$= (-1)^\alpha (m-\alpha)! n(t, U^{[\alpha]}, m-\alpha), \alpha = 0(1)q-d_p.$$

Es folgt mit (1.17) und (2.2):

$$\begin{aligned} & - \sum_{l=0}^m z^{-(l+1)} v_\alpha^{(1)}(0) \\ & = - \frac{(-z)^\alpha}{z^{m+1}} (m-\alpha)! \sum_{l=0}^{m-\alpha} z^{m-\alpha-l} n^{(1)}(0, U^{[\alpha]}, m-\alpha) \\ (+) & = - \frac{m!}{z^{m+1}} (-z)^\alpha \frac{(m-\alpha)!}{m!} \underbrace{\sum_{l=0}^{m-\alpha} z^l n^{(m-\alpha-l)}(0, U^{[\alpha]}, m-\alpha)}_{\tilde{q}(z, U^{[\alpha]}, m-\alpha)} \\ & = - \frac{m!}{z^{m+1}} q_\alpha(z), \alpha = 0(1)q-d_p. \end{aligned}$$

Analog zeigt man:

$$- \sum_{l=0}^m z^{-(l+1)} v_\alpha^{(1)}(1) = - \frac{m!}{z^{m+1}} p_\alpha(z), \alpha = 0(1)q-d_p.$$

Die Regularität der in (2.3) definierten Matrix H ist einfach einzusehen (vgl. S 6.3.a). Folglich ist obiges Gleichungssystem für $z \in \mathbb{C} - N_{q_0} - \{0\}$ nach χ_k auflösbar - die Lösung ist in $z=0$ stetig ergänzbar - und mit $\chi_0 = e_0$ folgt $\chi_k = \psi(z)^k e_0$, d.h.

$$(2.6) \quad \phi_k(t, z) = \frac{-z^{m+1}}{m! q_0(z)} v(t)^T \psi(z)^k e_0, t \in [0, 1], z \in \mathbb{C} - N_{q_0}, k \in \mathbb{N}_0.$$

Für die gesuchten Größen $s_k(z)$ folgt mittels (2.5), (***) und (*):

$$\begin{aligned} s_k(z) &= w_k(k) = \frac{z^{m+1}}{m! q_0(z)} \sum_{l=0}^m z^{-(l+1)} v^{(1)}(0)^T \psi(z)^k e_0 \\ &= \frac{1}{q_0(z)} q(z)^T \psi(z)^k e_0, \quad z \in \mathbb{C} - N_{q_0}, k \in \mathbb{N}_0 \end{aligned}$$

Korollar

Seien

$$D(z) := \begin{pmatrix} \frac{-m!}{z^{m+1}} p(z)^T \\ h^*, H^* \end{pmatrix}, \quad E(z) := \begin{pmatrix} \frac{-m!}{z^{m+1}} q(z)^T \\ 0, H \end{pmatrix},$$

$$T(z) := \begin{pmatrix} 1 & & & & \\ -z^{d_p+1} & 1 & & & \\ & -z & \ddots & & \\ & & \ddots & \ddots & \\ & & & -z & 1 \end{pmatrix} \in L(\mathbb{C}^{q-d_p+1}), z \in \mathbb{C} - \{0\}.$$

a) Es gilt: Die Größen $\varphi_k := \varphi_k(z) := (s_k(z), \phi_k^{(d_p)}(0, z), \dots, \phi_k^{(q-1)}(0, z))^T$ genügen der Rekursion

$$\varphi_{k+1} = D(z) E(z)^{-1} \varphi_k, \quad k \in \mathbb{N}_0, \quad \varphi_0 = e_0, \quad z \in \mathbb{C} - \mathbb{N}_{q_0} - \{0\}.$$

b) Eine Darstellung der Größen $s_k(z)$ mittels der Verfahrensparameter läßt sich auch aus der natürlichen matriziellen Beschreibung des Verfahrens (Vgl. Mülthei [28], S.452) herleiten. Nach Elimination der ersten $p-1$ Komponenten von S_k erhält man mit

$$\begin{aligned} S'_k &:= S'_k(z) := (s(kh), h^{d_p+1} s^{(d_p+1)}(kh), \dots, h^q s^{(q)}(kh))^T \\ S'_k &= \sum(z)^k S'_0, \text{ mithin } s_k(z) = e_0^T \sum(z)^k S'_0. \end{aligned}$$

Hierbei ist $\sum: \mathbb{C} - \mathbb{N}_{q_0} - \{0\} \longrightarrow L(\mathbb{C}^{q-d_p+1})$ gegeben durch:

$$\sum = T^{-1} D E^{-1} T = (E^{-1} T)^{-1} \Psi (E^{-1} T).$$

Bemerkung

Nachteil der unter b) angesprochenen Vorgehensweise ist, daß die Matrix $\sum(z)$ schon im einfachsten Fall $p=1$ in allen Komponenten z -abhängig ist, die Matrix $\Psi(z)$ dagegen nur in der nullten Zeile. Es sei aber vermerkt, daß auch die Übergangsmatrix der φ_k , nämlich $D(z) E(z)^{-1}$, "voll z -abhängig" ist. Der Trick, die Übergangsmatrix $\Psi(z) = E(z)^{-1} D(z)$ der χ_k zu betrachten, hat noch einen weiteren Effekt: Ψ ist stets in $z=\infty$ komponentenweise beschränkt - wie in Kapitel 4 gezeigt wird -, $D E^{-1}$ sowie $\sum$ haben diese Eigenschaft im allgemeinen nicht, was zu dem Fehlschluß der Nichtexistenz A -stabiler Verfahren in $V^>$ führen könnte, da A -Stabilität $\rho(\sum(\infty)) \leq 1$ erfordert. Dahingegen kann man an $\Psi(z)$ relativ leicht ablesen, daß im Fall der Konfluenz (Siehe Einleitung) ein Eigenwert von $\Psi(z)$ in der Nähe der Wachstumsfunktion des zugeordneten Verfahrens aus $V^<$, die anderen Eigenwerte in der Nähe von Null liegen.

Beweis

a) Die linke Seite des im Beweis zu T 2.2 aufgestellten Gleichungssystems für χ_k bedeutet

$$(*) \quad \frac{-z^{m+1}}{m!q_b(z)} E(z) \chi_k = \varphi_k.$$

Die Rekursion für die χ_k , nämlich $\chi_{k+1} = E(z)^{-1} D(z) \chi_k$, $k \in \mathbb{N}_0$, $\chi_0 = e_0$, impliziert mit (*) diejenige für die φ_k .
 b) Es gilt $\varphi_k = T(z) S_k^1$, $k \in \mathbb{N}_0$, $z \in \mathbb{C} - N_{q_b} - \{0\}$

Darstellung (2.6) besagt, daß die Polynome $\phi_k(.,z)$, $k \in \mathbb{N}_0$, neben $m-(q-d_p)$ festen Nullstellen U_2 (Vgl. D 1.1.d) noch (höchstens) $q-d_p$ weitere freie, von z - und k -abhängige Nullstellen besitzen. Letztere sowie die Koeffizienten eines Polynoms, welches diese als Nullstellen besitzt und die Funktionen R_k aus (1.12) sollen im folgenden definiert und mittels Ψ dargestellt werden. Darauf wird die Konvergenz dieser Größen untersucht, was schließlich die gewünschte Darstellung (1.13) der Funktion R aus (1.12) ergibt.

2.3 Definition

a) Sei

$$R_k(z) := \begin{cases} \frac{s_{k+1}(z)}{s_k(z)} & , s_k(z) \neq 0, \\ \infty & , s_k(z) = 0, s_{k+1}(z) \neq 0, \\ 0 & , s_k(z) = s_{k+1}(z) = 0, z \in \mathbb{C} - N_{q_b}, k \in \mathbb{N}_0. \end{cases} \quad 1)$$

b) Seien $p_0^C := \tilde{p}(., (\underbrace{0, \dots, 0}_{d_p}, \underbrace{T, 1, \dots, 1}_q), m)$, $z \in \mathbb{C} - N_{q_b} - N_{p_0^C}$ und $k \in \mathbb{N}_0$.

Weiter seien

$$\mu_k := \mu_k(z) := \min \{ i \in \{0, \dots, q-d_p\} \mid e_i^T \Psi(z)^k e_0 \neq 0 \}, \quad 2)$$

¹⁾ In einigen Punkten z ist also das Produkt $R_{k-1}(z) \cdot \dots \cdot R_0(z)$ aus (1.11) durch die Größe $s_k(z)$ definiert.

²⁾ Aufgrund der Wahl von z ist $\Psi(z)$ regulär, weswegen $\mu_k \in \{0, \dots, q-d_p\}$ existiert (Vgl. S 6.3.b(ii)).

$j_k := j_k(z) := q - d_p - \mu_k$ und ferner der Vektor

$$U_1^{[k]}(z) := (u_1^{[k]}(z), \dots, u_{j_k}^{[k]}(z)) \in \mathbb{C}^{j_k}$$

dadurch definiert, daß $(U_1^{[k]}(z), U_2^{[k]}(z)) \in \mathbb{C}^{m-\mu_k}$ die Nullstellen des Polynoms $v(\cdot)^T \Psi(z)^k e_0 \in \prod_{m-\mu_k}$ bezeichne. Zuletzt sei

$$x^{[k]}(z) := \begin{pmatrix} e_{\mu_k}^T \\ \vdots \\ e_{q-d_p}^T \end{pmatrix} \Psi(z)^k e_0 / e_{\mu_k}^T \Psi(z)^k e_0$$

der Koeffizientenvektor desjenigen Polynoms mit Hauptkoeffizient 1, welches die Nullstellen $U_1^{[k]}(z)$ besitzt¹⁾.

Bemerkung

Aus (2.5) und (2.6) ist die Bedeutung der Zahlen $u_1^{[k]}(z), \dots, u_{j_k}^{[k]}(z)$ ersichtlich: Offensichtlich gilt neben

$$(2.7) \quad \begin{cases} s_k^{(j)}((k+t_i)h, h) = \mu s_k^{(j-1)}((k+t_i)h, h), \\ i=0(1)p, j=1(1)d_i^{(2)}, \text{ auch} \\ s_k^{(j)}((k+u_j^{[k]}(z))h, h) = \mu s_k^{(j-1)}((k+u_j^{[k]}(z))h, h), j=1(1)j_k. \end{cases}$$

Man kann also sagen, daß die Zahlen $u_j^{[k]}(z)$, $j=1(1)j_k$, (bzgl. (T_μ)) freie Kollokationsstellen sind, die sich das Verfahren in jedem Schritt "selbst zurechtlegt". Die Beschreibung hängt in der Tat nur von den symmetrischen Funktionen dieser freien Kollokationsstellen ab, dasselbe gilt jedoch auch für die festen Zwischenknoten. Andererseits sind die $U_1^{[k]}(z)$ im allgemeinen nicht explizit gegeben, was jedoch weder für die angestrebten Resultate noch für mögliche Interpretationen eine Rolle spielt.

¹⁾ Es gilt

$$\prod_{j=1}^{j_k} (t - u_j^{[k]}(z)) = \sum_{j=0}^{j_k} t^{j_k-j} (-1)^j x_j^{[k]}(z), \quad t \in \mathbb{C}.$$

²⁾ $d_0' := d_p$; $d_i' := d_i$, $i=1(1)p$.

2.4 Satz

a) Für die in D 2.3 definierten Funktionen R_k gelten die beiden folgenden Darstellungen:

$$(2.8) \quad R_k(z) = \frac{q(z)^T \psi(z)^{k+1} e_0}{q(z)^T \psi(z)^k e_0}, \quad z \in \mathbb{C} - N_{q_0},$$

$$(2.9) \quad R_k(z) = \tilde{Q}(z, (U_1^{[k]}(z), U_2), m - \mu_k), \quad z \in \mathbb{C} - N_{q_0} - N_{p_0}, \quad k \in \mathbb{N}_0.$$

b) Die symmetrischen Funktionen der freien Kollokationsstellen unterliegen für $z \in \mathbb{C} - N_{q_0} - N_{p_0}$ der Rekursionsvorschrift

$$(2.10) \quad \left\{ \begin{array}{l} x^{[0]}(z) = e_0, \\ x^{[k+1]}(z) = \frac{\psi(z)^{k+1} e_0}{e_0^T \psi(z)^{k+1} e_0} = \frac{\psi(z) x^{[k]}(z)}{e_0^T \psi(z) x^{[k]}(z)}, \\ \text{falls } e_0^T \psi(z)^j e_0 \neq 0, \quad j=k, k+1, \quad k \in \mathbb{N}_0. \end{array} \right.$$

Bemerkungen

a) Formel (2.9) besagt, daß die R_k , $k \in \mathbb{N}$, formal die gleiche Struktur besitzen wie R_0 (Vgl. T 1.2.b). Im allgemeinen wird $\mu_k = 0$, $k \gg k_0$, sein, woran man schon sehen kann, wie die gesuchte Darstellung (1.13) zustande kommt. Gemäß der Bemerkung zu L 1.3 läßt sich sogar (formal)

$$R_k(z) = \tilde{Q}(z, (\underbrace{U_1^{[k]}(z), \infty, \dots, \infty}_{\mu_k}, U_2), m)$$

schreiben, womit man sagen kann, daß die Vielfachheiten Nummer d_0+1 bis q zu $t_0 = 0$ gegenüber R_0 bei R_k durch genau $q-d_0$ z - und k -abhängige freie Kollokationsstellen in $\mathbb{K}$ ersetzt sind.

b) Die Formeln (2.10) und (2.8) besagen, daß es sich bei $(x^{[k]}(z), R_k(z))$ um eine Eigenvektor-Eigenwert-Iteration bzgl. $\psi(z)$ handelt (Genaueres im folgenden Abschnitt 2-2).

Beweis

a) Formel (2.8) folgt unmittelbar aus D 2.3.a mit (2.1). Aus der Verallgemeinerung des Beweises zu Theorem 4 aus Nörsett/Wanner

[32] ergibt sich mittels (1.17) die Darstellung (2.9) von $R_k(z)$, wenn man berücksichtigt, daß dort einerseits nur eingeht, daß $\phi_0(.,z)$ den Höchstgrad m besitzt und es andererseits keine Rolle spielt, ob die Kollokationspunkte reell oder komplex sind. Für $z \in \mathbb{C} - N_{q_0} - N_{p_0} \subset \mathbb{C}$ sind alle Größen in (2.9) definiert, wobei $R_k(z) = \infty$ möglich ist und im Fall $\tilde{p}(z, (u_1^{[k]}(z), u_2), m - \mu_k) = \tilde{q}(z, (u_1^{[k]}(z), u_2), m - \mu_k) = 0$ entsprechend $D \ 2.3.a \ R_k(z) := 0$ sei.

- b) Der Einfachheit halber werde hier das Argument z weggelassen. Sei für $k \in \mathbb{N}_0$ $e_0^T \psi^j e_0 \neq 0$, $j = k, k+1$. Die beiden ersten Gleichungen folgen unmittelbar aus $D \ 2.3.b$, die dritte erhält man, indem man $\psi^{k+1} e_0 / e_0^T \psi^{k+1} e_0$ mit $e_0^T \psi^k e_0$ erweitert.

Im folgenden Korollar wird (2.10) noch einmal speziell für $q = d_p + 1$ angegeben, da für die angestrebte Charakterisierung der Konvergenzordnung in gewissen -hier nur für $q = d_p + 1$ untersuchten- Randfällen Laurententwicklungen der Fixpunkte von (2.10) um $z=0$ benötigt werden.

2.5 Korollar

Sei $q = d_p + 1$. Mit den Polynomen $a := -q_0$, $c := \beta p_1 + q_1$, $d := \beta p_0 - q_1$, mit

$$(2.11) \quad \beta := (-1)^{m-d} p \prod_{i=1}^{p-1} \left(\frac{t_i}{1-t_i} \right)^{d_i} \quad \text{und}$$

$$u_k(z) := \begin{cases} u_1^{[k]}(z) & , j_k = 1, \\ \infty & , j_k = 0, \end{cases} \quad z \in \mathbb{C} - N_{q_0} - N_{p_0} \subset \mathbb{C},$$

gilt für die (eine) freie Kollokationsstelle und $z \in \mathbb{C} - N_{q_0} - N_{p_0} \subset \mathbb{C}$ stets die Rekursionsvorschrift

$$(2.12) \quad u_0(z) = 0, \quad u_{k+1}(z) = \varphi(u_k(z), z) := \frac{a(z)(u_k(z) - 1)}{c(z)u_k(z) + d(z)}, \quad k \in \mathbb{N}_0.$$

Beweis

Die Gleichung für $\varphi(u, z)$ folgt aus (2.10) und (2.2) mit $x^{[k]}(z) = (1, u_1^{[k]}(z))^T$ sowie $H^{-1}h^* = \beta^{-1}$ und $H^{-1}H^* = -\beta^{-1}$ (Siehe auch Beispiel a zu S 6.3). Zudem muß man sich noch überlegen, daß im Fall $e_0^T \psi^k e_0 = 0$ ($u_k = \infty$), der nur für $c \neq 0$ auftreten kann, das durch die Vorschrift (2.12) bestimmte $u_{k+1} = a/c$ mit dem vom Verfahren definierten $u_1^{[k+1]}$ übereinstimmt, was hier nicht ausgeführt werden soll.

2-2 Konvergenz der freien Kollokationsstellen

In diesem Abschnitt soll die Konvergenz der freien Kollokationsstellen untersucht werden. Zu diesem Zweck sind die Größen $(x^{[k]}(z), R_k(z))$, $k \in \mathbb{N}_0$, zu betrachten. Es handelt sich hierbei, wie gesagt, um eine Eigenvektor-Eigenwert-Iteration. Die Eigenwerte λ der Matrix $\Psi^{(1)}$ genügen einer algebraischen Gleichung der Ordnung $q-d_p+1$, die mit

$$(2.13) \quad f_{\Psi(z)}(\lambda) = \lambda^{q-d_p+1} + a_{q-d_p}(z) \lambda^{q-d_p} + \dots + a_0(z) = 0$$

bezeichnet werde. Im folgenden sollen der Einfachheit halber nurmehr Verfahren betrachtet werden, für die

$$(2.14) \quad \text{ggT}(q_0, q_0 a_{q-d_p}, \dots, q_0 a_0) = 1^{(2)}$$

gilt. Diese Bedingung wird benötigt für die Charakterisierung der Region absoluter Stabilität durch die noch zu definierende Funktion R ((2.16)) sowie für gewisse Holomorphieaussagen, nicht jedoch für die Charakterisierung der Konvergenzordnung in Kapitel 3. Sie bedeutet ganz einfach, daß für $z \in \mathbb{N}_{q_0}$ mindestens einer der Eigenwerte von $\Psi^{(1)}$ einen Pol besitzt. Zu $z \in \mathbb{K} \setminus \mathbb{N}_{q_0}^{(3)}$ bezeichne

$$(\lambda_0(z), \dots, \lambda_{q-d_p}(z))$$

das $(q-d_p+1)$ -Tupel der dem Betrage nach fallend angeordneten Eigenwerte von $\Psi(z)$ (Sofern mehrere betragsgleiche Eigenwerte existieren, sei hiermit eine beliebige, aber feste Reihenfolge fixiert). Ferner sei gemäß obiger Bemerkung

$$\lambda_0(z) := \infty, z \in \mathbb{N}_{q_0}.$$

Gemäß Bemerkung c zu T 2.2 kann man sich die Werte $(s_k(z))_{k \in \mathbb{N}_0}$ auch durch eine lineare Differenzengleichung erzeugt denken. Mithin genügen bekanntlich die Vektoren

$$(s_k(z), \dots, s_{k+q-d_p}(z))^T, k \in \mathbb{N}_0,$$

2) Aus (2.2) bzw. (6.2) folgt: $a_i = \tilde{a}_i / q_0$ mit Polynomen $\tilde{a}_i$, $i=0(1)q-d_p$. Unter $q_0 a_i$ werde $\tilde{a}_i$ verstanden.

1) Eigentlich: Die Eigenwertfunktionen der matrixwertigen Fkt. Ψ .
 2) Bezüglich $z = \infty$ siehe S 4.2.

einer Vektoriteration (I) (Vgl. z.B. Stoer/Bulirsch [41], 7.2.9). Der Index $(:= \nu(z))$ sei dadurch charakterisiert, daß der Startvektor zu (I) nicht an den Eigen- und Hauptvektoren der Frobeniusform von Ψ zu $\lambda_0, \dots, \lambda_{l-1}$, wohl aber an mindestens einem der Eigen- und Hauptvektoren der Eigenwerte vom "nächstkleineren" Betrage $|\lambda_l|$ ($|\lambda_{l-1}| > |\lambda_l|$ (1)) beteiligt ist.

Damit kann schließlich die Funktion R definiert werden:

2.6 Definition

$$(2.15) \quad R(z) := \begin{cases} \lim_{k \rightarrow \infty} R_k(z) & , z \in \mathbb{C} - N_{q_0} - A, \\ \infty & , z \in N_{q_0}, \\ \lambda_l(z) & , z \in A, \\ \lambda_0(\infty) & , z = \infty, \end{cases} \quad \text{mit}$$

$$A := \{ z \in \mathbb{C} - N_{q_0} \mid \lim_{k \rightarrow \infty} R_k(z) \text{ existiert nicht} \}.$$

Bemerkungen

- Die Definition, die R auf ganz $\mathbb{K}$ festlegt, beinhaltet als Grundgedanken $R := \lim_{k \rightarrow \infty} R_k$. Diese Idee wird über Formel (2.9) zur Darstellung (1.13) und zur Charakterisierung der Konvergenzordnung führen.
- Da es sich bei R stets um einen Eigenwert von Ψ handelt, könnte man von "verallgemeinertem Eigenwert" sprechen. Dieser bestimmt für $z \in \mathbb{C} - N_{q_0} - A$ das Wachstumsverhalten der $s_k(z)$ für hinreichend großes k:

$$s_{k+1}(z) \approx R(z)s_k(z), \quad k \gg k_0(z), \quad z \in \mathbb{C} - N_{q_0} - A$$

(Vgl. D 2.3 a). Man überlegt sich mittels Formel (7.2.9.20) aus Stoer/Bulirsch [41], daß $|R(z)| = |\lambda_l(z)|$, $z \in \mathbb{C} - N_{q_0}$, gilt, was wegen der Aussage " $\lim_{k \rightarrow \infty} s_k(z) = 0 \Leftrightarrow |\lambda_l(z)| < 1$ " mit D 2.1 zur Verallgemeinerung von (1.10) auf den Fall $q > d_p$ führt:

$$(2.16) \quad \begin{aligned} R_A &= \{ z \in \mathbb{C} \mid |R(z)| < 1 \} \\ &= \{ z \in \mathbb{C} \mid q(\Psi(z)) < 1 \} \cup \\ &\quad \{ z \in \mathbb{C} \mid |\lambda_l(z)| < 1, |\lambda_0(z)| \geq 1 \} \end{aligned}$$

Um den angestrebten Konvergenzsatz formulieren zu können, ist die Definition dreier Ausnahmemengen erforderlich:

2.7 Definition

$$\begin{aligned} \text{Seien} \quad A' &:= \{ z \in \mathbb{C} - N_{q_0} \mid |\lambda_0(z)| = |\lambda_1(z)| \}, \\ B &:= \{ z \in \mathbb{C} - N_{q_0} \mid |\lambda_0(z)| > |\lambda_1(z)| \}, \\ C &:= \{ z \in \mathbb{C} - N_{q_0} - N_{p_0}^{C-A'-B} \mid \lim_{k \rightarrow \infty} x_r^{[k]}(z)^{1)} \text{ existiert} \\ &\quad \text{nicht in } \mathbb{C}^{q-d_p} \}, \\ F &:= N_{q_0} \cup N_{p_0}^{C-A'-B} \cup \{\infty\} \cup A' \cup B \cup C. \end{aligned}$$

Bemerkungen

- a) Die Ausnahmemengen in $\mathbb{K}$, nämlich N_{q_0} und $\{\infty\}$, $N_{p_0}^C$ (und $\{0\}$) sowie A', B, C sind begrifflich einfach zu verstehen:

In N_{q_0} und $\{\infty\}$ ist das Verfahren (bzgl. (T_μ)) undefiniert. In $N_{p_0}^C$ (und $\{0\}$) sind die freien Kollokationsstellen im allgemeinen nicht definierbar (, in $\{0\}$ stets stetig ergänzbar). Für $z \notin A' \cup B$ ist die Eigenwertiteration $(R_k(z))_{k \in \mathbb{N}_0}$ -vergleiche (2.8)- am betragsgrößten Eigenwert von $\Psi(z)$ beteiligt und dieser ist einfach. Für $z \notin C$ konvergiert auch die Eigenvektoriteration $(x^{[k]}(z))_{k \in \mathbb{N}_0}$, mit anderen Worten besitzt $s_k^1(\cdot, h) - \mu s_k(\cdot, h)$ nur endlich oft weniger als $q-d_p$ freie Kollokationsstellen und letztere konvergieren (bei geeigneter Anordnung) gegen $q-d_p$ Grenzwerte

$$u_1(z), \dots, u_{q-d_p}(z) \in \mathbb{C},$$

die implizite Kollokationsstellen heißen sollen (Genauer: I 2.8).

- b) Die Charakterisierung der Mengen A (bzw. A'), B und C durch die dem Verfahren zugeordneten Größen ist jedoch kompliziert, sie findet sich im Anhang (S 6.4) für den Fall $q = d_p + 1$.
- c) Für die Charakterisierung der Konvergenzordnung sowie für die Untersuchung auf A-Stabilität wird es wichtig sein zu wissen, wo diese Mengen liegen. Genauer benötigt man zu gegebenem $\varepsilon > 0$ für ersteres $F \cap (0, \varepsilon] = \emptyset$ und bezüglich A-Stabilität $F \cap [-i\varepsilon, i\varepsilon] = \emptyset$, $A' \cap i\mathbb{R} = \emptyset$, $N_{q_0} \cap \mathbb{C}^- = \emptyset$.

1) Zu $x^{[k]}(z) = (x_0^{[k]}(z), \dots, x_{j_k}^{[k]}(z))^T$ sei $x_r^{[k]}(z) := (x_1^{[k]}(z), \dots, x_{j_k}^{[k]}(z))^T$.

2.8 Theorem

Sei $z \in \mathbb{K} - F$. Dann konvergieren die Vektoren $x^{[k]}(z)$ gegen den auf $x_0(z) = 1$ normierten Eigenvektor $x(z)$ zum Eigenwert $R(z)$ von $V(z)$. In Formeln:

$$(V(z) - R(z)I) x(z) = 0, \quad x(z) := \lim_{k \rightarrow \infty} x^{[k]}(z).$$

Bemerkungen

- a) Es sei noch einmal daran erinnert, daß die $x^{[k]}(z)$ gerade die Koeffizientenvektoren¹⁾ der Polynome $h_k(., z)$ sind, die den freien Kollokationsstellen vermöge

$$h_k(t, z) := \prod_{j=1}^{j_k} (t - u_j^{[k]}(z)), \quad t \in \mathbb{R},$$

zugeordnet sind. Der Gehalt der Behauptung besteht also darin, daß die freien Kollokationsstellen $u_1^{[k]}(z), \dots, u_{j_k}^{[k]}(z)$, $k \in \mathbb{N}_0$, (bei geeigneter Anordnung) gegen $q-d_0$ Grenzwerte $u_1(z), \dots, u_{q-d_0}(z)$ in $\mathbb{C}$ konvergieren und letztere (in Form ihrer symmetrischen Grundfunktionen Nummer 0 bis $q-d_0$) einen Eigenvektor zu $R(z)$ von $V(z)$ bilden.

- b) Die Bedeutung dieser Grenzwerte, die (ja) implizite Kollokationsstellen heißen sollen, liegt gemäß (2.7) darin, daß das Verfahren diese nach "genügend langer Rechnung" tatsächlich approximativ als Kollokationsstellen besitzt:

$$s_k'((k+u_j(z))h, h) \approx \mu s_k'((k+u_j(z))h, h), \quad j=1(1)q-d_0, k \neq k_0(z).$$

Bei festem $z := h\mu \in \mathbb{K} - F$ verhält sich das Verfahren "weit draußen" also, als besäße es die Kollokationsstellen

$$(2.17) \quad \begin{cases} U(z) &:= (u_1(z), u_2(z)), \\ U_1(z) &:= (u_1(z), \dots, u_{q-d_0}(z)). \end{cases}$$

Im anschließenden Kapitel über die Charakterisierung der Konvergenzordnung wird die anschauliche Bedeutung dieser Tatsache demonstriert.

1) Vergleiche Fußnote 1) zu D 2.3.b .

Korollar

Seien $z \in \mathbb{K} - F$ und $R(z)$ kein Eigenwert von $H^{-1}H^*$. Dann folgt die Identität

$$(2.18) \quad x_I(z) = (HR(z) - H^*)^{-1}h^*.$$

Beweis des Korollars

Aus (2.2) bzw. dem Beweis zu S 6.3.a ergibt sich mit T 2.8

$$\begin{pmatrix} e_1^T \\ \vdots \\ e_{q-d_p}^T \end{pmatrix} \Psi(z)x(z) = H^{-1}(h^*, H^*)x(z) = R(z)x_I(z),$$

also folgt $(HR(z) - H^*)x_I(z) = h^*$.

Beweis des Theorems

Sei $z \in \mathbb{K} - F$. Der Einfachheit halber werde im folgenden das Argument z weggelassen. Wegen $z \in \mathbb{C} - N_{Q_0} - A'$ folgt

(*) λ_0 ist einfacher Eigenwert mit $\alpha := |\lambda_1|/|\lambda_0| < 1$.

Da sogar $z \in \mathbb{C} - N_{Q_0} - (A' \cup B)$ ist, ergibt sich

(**) Die Folge $(R_k)_{k \in \mathbb{N}_0}$ konvergiert gegen $R = \lambda_0$.

Wegen $z \notin C$ existiert der Grenzwert $x := \lim_{k \rightarrow \infty} x^{[k]}$ in $\mathbb{C}^{q-d_p+1}$. Sei y_0 Eigenvektor zu λ_0 , ergänzt zu einem System $\{y_0, \dots, y_{q-d_p}\}$ von Eigen- und Hauptvektoren der Matrix Ψ sowie

$$(***) \quad e_0 = \alpha_0 y_0 + \sum_{i=1}^{q-d_p} \alpha_i y_i$$

die eindeutig bestimmte Zerlegung des Startvektors von (2.10) in die Vektoren dieses linear unabhängigen Systems.

Zu zeigen sind noch $e_0^T y_0 \neq 0$ und $x = y_0 / e_0^T y_0$. Zunächst folgen $\alpha_0 \neq 0$ sowie $e_0^T y_0 \neq 0$, da die Folge der R_k wegen (2.8) und (*) andernfalls entweder nicht oder gegen einen Eigenwert von geringerem Betrag als $|\lambda_0|$ konvergierte, was (**) widerspräche. Mit (***) folgt

$$\Psi^k e_0 = \alpha_0 \lambda_0^k \left(y_0 + \sum_{i=1}^{q-d_p} \alpha_i \lambda_i^k y_i \right).$$

Wegen $z \notin C$, D 2.7 und D 2.3.b existiert ein $k_0 \in \mathbb{N}$ mit:

$$x^{[k]} = \frac{\psi^k e_0}{e_0^T \psi^k e_0} = \frac{y_0 + O(\alpha^k)}{e_0^T y_0 + e_0^T O(\alpha^k)}, \quad k \geq k_0.$$

Die Konvergenz der Folge $(x^{[k]})_{k \in \mathbb{N}_0}$ liefert $e_0^T y_0 \neq 0$ sowie $x = y_0 / e_0^T y_0$.

Für das Folgende sei noch, wie soeben bewiesen,

$$(2.19) \quad Q(z) := q(z)^T x(z) \neq 0, \quad z \in \mathbb{K} - F,$$

vermerkt.

2-3 Darstellungen und Stetigkeits-
eigenschaften von R

2.9 Definition

Sei $M \subset \mathbb{K}$ offen. $\mathcal{H}(M)$ bezeichne die Menge aller holomorphen (= lokal analytischen) Funktionen $f : M \rightarrow \mathbb{C}$.

Im folgenden Satz werden zwei Eigenschaften von R aufgezeigt, welche die wesentlichen Hilfsmittel zur Charakterisierung der Konvergenzordnung und Bestimmung A -stabiler Verfahren darstellen. Stetigkeit bzw. Holomorphie von $|\lambda_0(\cdot)|$ bzw. λ_0 gehen in Kapitel 4 und 5 ein.

2.10 Satz

- a) Die in D 2.6 definierte Funktion R ist bis auf eine Ausnahmemenge gleich dem betragsgrößten Eigenwert λ_0 der Matrix Ψ^1 :

$$(2.20) \quad \begin{aligned} R(z) &= \lambda_0(z), \quad z \in \mathbb{K} - A' - B, \\ |R(z)| &= |\lambda_0(z)|, \quad z \in \mathbb{K} - B, \\ |R(z)| &< |\lambda_0(z)|, \quad z \in B. \end{aligned}$$

- b) Die impliziten Kollokationsstellen aus (2.17) erlauben eine Darstellung von R mittels der Nörsett/Wannerschen Funktion $\tilde{Q}$ (Vergleiche T 1.2) :

$$(2.21) \quad R(z) = \tilde{Q}(z, U(z), m) = \frac{p(z)^T x(z)}{q(z)^T x(z)}, \quad q(z)^T x(z) \neq 0, \quad z \in \mathbb{K} - F.$$

Weiter gilt im Fall $x_{q-d_p}(z) \neq 0$ ²⁾ mit β aus (2.11)

$$(2.22) \quad R(z) = \frac{1}{\beta} \prod_{j=1}^{q-d_p} \frac{1 - u_j(z)}{u_j(z)}, \quad z \in \mathbb{K} - F.$$

- c) Es gilt $|\lambda_0(\cdot)| \in C^0(\mathbb{K} - N_{q_0})$.
 d) Unter den zusätzlichen Voraussetzungen $ggT(q_0, q_0 a_{q-d_p}) = ggT(q_0 a_{q-d_p}, \dots, q_0 a_0) = 1$ ³⁾ an die in (2.13) definierten

¹⁾ Vergleiche Fußnote 1) zur Einleitung des Abschnittes 2-2.

²⁾ Im Fall $q = d_p + 1$ gilt $x_{q-d_p}(z) = u_1(z) \neq 0, \quad z \in \mathbb{K} - F$.

³⁾ Im Fall $q = d_p + 1: N_{q_0} \cap N_{a+d} = N_{a+d} \cap N_{p_0^c} = \emptyset$ (Vgl. K 2.5).

Funktionen a_i , $i=0(1)q-d_p$, folgt

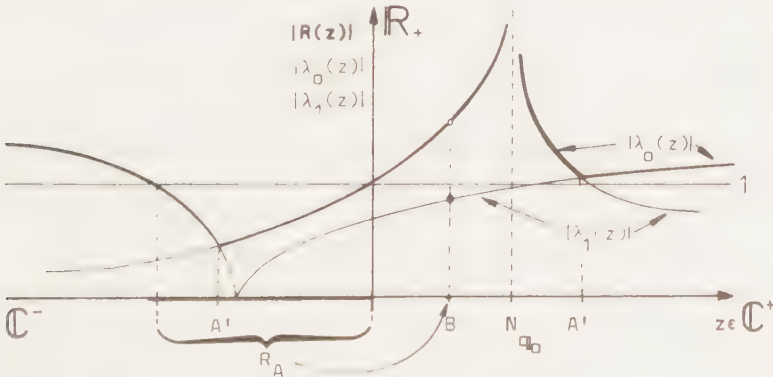
$$\lambda_0 \in \begin{cases} \mathcal{X}(\mathbb{K}-N_{q_0}-A') & , |\lambda_0(\infty)| > |\lambda_1(\infty)| , \\ \mathcal{X}(\mathbb{C}-N_{q_0}-A') & , |\lambda_0(\infty)| = |\lambda_1(\infty)| . \end{cases}$$

Bemerkungen

- a) Die Darstellung (2.21) sowie L 6.1.f zeigen, daß die Reihenfolge der impliziten Kollokationsstellen $u_j(z)$, $j=1(1)q-d_p$, nicht eingeht. Formel (2.22), die für das Folgende nicht benötigt wird, ist von daher interessant, daß sie mit Ausnahme von $q=d_p$ und $z=\infty$ keine Entsprechung in V^* hat.
- b) Aus S 2.10.a und c folgen unmittelbar

$$|R(\cdot)| \in C^0(\mathbb{K}-N_{q_0}-\bar{B}), R \in \mathcal{X}(\mathbb{C}-N_{q_0}-A'-\bar{B})^{1)}.$$

Die folgende Skizze möge die Stetigkeits- und Differenzierbarkeitsverhältnisse von R demonstrieren:



Beweis

- a) Die Gleichheit $R(z) = \lambda_0(z)$ folgt für $z = \infty$ und $z \in N_{q_0}$ aus D 2.6, für $z \in \mathbb{C}-N_{q_0}-(A' \cup B)$ aus dem Beweis zu T 2.8. Die bereits in Bemerkung b zu D 2.6 angedeutete Aussage

$$|R(z)| = |\lambda_l(z)|, z \in \mathbb{C}-N_{q_0} \supset (A' \cup B),$$

impliziert mit D 2.7 die beiden übrigen Behauptungen.

1) Unter dem Abschluß $\bar{M}$ einer Menge M werde stets der Abschluß in $\mathbb{K}$ verstanden.

- b) Das erste Gleichheitszeichen in (2.21) folgt unmittelbar aus Formel (2.9) und T 2.8, weiter gilt wegen T 2.8 und (2.19) für $z \in \mathbb{K} - F : \psi x = Rx, \quad q_1^T x \neq 0$. Es folgt mittels S 6.3.b(i)

$$R = q_1^T \psi x / q_1^T x = p^T x / q_1^T x.$$

Der Beweis des Korollars zu T 2.8 besagt $RHx_r = h^* + H^* x_r$. Formel (2.22) erhält man, indem man auf beiden Seiten dieser Gleichung zwischen Vektoren jeweils die erste Komponente betrachtet (Siehe Anhang, Beweis zu S 6.3.a, (6.3) und (6.1)).

- c) Die Koeffizienten von ψ hängen in $\mathbb{K} - N_{q_0}$ stetig von z ab, was für den Punkt $z = \infty$ in Kapitel 4 gezeigt wird. Die Identität $|\lambda_0(\cdot)| = g(\psi(\cdot))$ ergibt (bekanntlich) die Behauptung.
- d) Es wird nur $|\lambda_0(\infty)| > |\lambda_1(\infty)|$ betrachtet, der andere Fall ist analog zu behandeln. Neben der Aussage in c gilt bekanntlich ebenso $|\lambda_1(\cdot)| \in C^0(\mathbb{K} - N_{q_0})$. Aus den beiden ggT-Bedingungen in der Voraussetzung folgt mit (2.13) einerseits, daß $|\lambda_1(\cdot)|$ in N_{q_0} stetig fortsetzbar ist, andererseits, daß λ_0 auf ganz $\mathbb{K}$ nicht verschwindet. Folglich ist die Abbildung

$$g : z \mapsto |\lambda_1(z)/\lambda_0(z)|$$

stetig in $\mathbb{K}$ mit $|g(z)| \leq 1, z \in \mathbb{K}$ und damit $\mathbb{K} - A' - N_{q_0} = g^{-1}(B_1(0)) - N_{q_0}$ offene Teilmenge von $\mathbb{K} \setminus (B_1(0) := \{z \in \mathbb{C} \mid |z| < 1\})$.

Die durch die algebraische Gleichung (2.13) definierte algebraische Funktion λ ist auf der zugeordneten Riemannschen Fläche $\tilde{X}_\lambda$ holomorph bis auf ihre Pole N_{q_0} . Es folgt $\lambda_0 \in \mathcal{H}(\mathbb{K} - N_{q_0} - A')$.

(Hierbei ist die gesamte Menge A' "herausgenommen", da dort, also in A' , λ_0 nicht in dem hier benötigten Sinne¹⁾ eindeutig bestimmt, sondern nur willkürlich festgelegt ist. Innerhalb der einzelnen Zusammenhangskomponenten von $\mathbb{K} - N_{q_0} - A'$ gilt sogar die eindeutige analytische Fortsetzbarkeit von jeder Einschränkung von λ_0 , die dort als Funktionskeim betrachtet werden kann.)

¹⁾ nämlich als Zweig der Funktion λ

3 Ordnung von R und V

3-1 Charakterisierung der globalen Konvergenzordnung von V

3.1 Definition

Sei $U^{[0]} \in V$.

- a) Eine Zahl $\nu \in \mathbb{Z}$ heie (globale) Konvergenzordnung (von V) oder Ordnung (von V), falls

$$(3.1) \quad \max_{x \in [0,1]} |y(x) - s(x,h)| = c_1(y) h^\nu + O(h^{\nu+1}), \quad h \rightarrow 0,$$

gilt. Dabei ist y die Lsung eines Anfangswertproblems der Form (1.1) und $s(\cdot, h)$ der von $U^{[0]} \in V$ zu (1.1) erzeugte Spline. Die Relation (3.1) hat fr smtliche f aus einer gewissen Stetigkeitsklasse bzw. die zugeordneten $y, s(\cdot, h)$ zu gelten (Vergleiche Mlthei [26], S.319), wobei c_1 ein auf dieser Klasse von Funktionen nicht verschwindendes Funktional ist.

- b) Eine Zahl $\bar{\nu} \in \mathbb{Z}$ heie lokale Konvergenzordnung (von V) oder (Konvergenz)ordnung (von V) in den Hauptknoten, falls mit $G_h := \{jh \mid j=0(1)n\}$, $h := 1/n$, $n \in \mathbb{N}$, fr jedes feste $x_0 \in (0,1)$ die Entwicklung

$$(3.2) \quad \max_{x \in G_h \cap [x_0, 1]} |y(x) - s(x,h)| = \bar{c}_1(y) h^{\bar{\nu}} + O(h^{\bar{\nu}+1}), \quad h \rightarrow 0,$$

gilt. Das unter a weiter Gesagte gelte entsprechend ($\bar{c}_1(y) = \bar{c}_1(y, x_0)$).

- c) Sei $\mathbb{R}_+ := \{z \in \mathbb{R} \mid z > 0\}$. Eine Zahl $\mu \in \mathbb{Z}$ heie Ordnung von R, falls

$$(3.3) \quad e^z - R(z) = c_2 z^{\mu+1} + O(z^{\mu+2}), \quad z \rightarrow 0, \quad z \in \mathbb{R}_+,$$

mit $c_2 \neq 0$ gilt. Von "Ordnung" $\mu = -1$ werde auch dann gesprochen, wenn zwar keine Entwicklung der Form (3.3) bekannt ist, jedoch der Grenzwert $\lim_{\substack{z \rightarrow 0 \\ z \neq 0}} |R(z)|$ existiert und verschieden von 1 ist.

Bemerkung

Fr jedes feste $x \in (0,1)$ ist der Grenzübergang fr $h \rightarrow 0$ in b so zu verstehen, da nur Schrittweiten h in Frage kommen, fr die $x \in G_h$ gilt.

Es sei ab sofort $m \geq 2$, was nur der Deutlichkeit halber gesagt werde, im Grunde aber keine Einschrnkung bedeutet.

Ziel dieses Abschnittes ist es, die Ordnung von V durch die von R zu charakterisieren.

Für die Konvergenz bzw. Divergenz eines Verfahrens $U^{[0]} \in V^>$ ist, in der Beschreibung von Mülthei, die Matrix $A_2 \in L(\mathbb{R}^{q-dp})$ verantwortlich, wobei (i.wes.) genau für die konvergenten Verfahren $g(A_2) \leq 1$ gilt (Vgl. [26], Satz 4, Satz 5, [27], Satz 2, [28], Satz 4). Den Verfahren aus verschiedenen Klassen, z.B. aus der Klasse der IRK-Methoden, der linearen MSV, der one-leg Methoden, PD-Verfahren, zusammengesetzten MSV, zyklischen Methoden oder MSV mit höheren Ableitungen (Aufzählung gemäß Jeltsch/Nevalinna [21]) läßt sich bei Anwendung auf (T_μ) , $z := h\mu$, eine charakteristische Relation der Form

$$(3.4) \quad \phi(\lambda, z) = \sum_{j=0}^M \sum_{i=0}^K a_{ij} z^j \lambda^i = 0$$

mit natürlichen Zahlen M, K und reellen Koeffizienten a_{ij} , $i=0(1)K$, $j=0(1)M$, zuordnen, was hier für die Verfahren aus der Klasse $V^>$ mit T 2.2 und (2.4) gezeigt wurde. Für lineare MSV und lineare MSV mit höheren Ableitungen, die dem Dahlquistischen Wurzelkriterium genügen, ist die Aussage " $\bar{v}$ = algebraische Ordnung des (sogenannten) Hauptzweiges von (3.4) " bekannt (Siehe Jeltsch/Nevalinna [20], S.63 und [21], S.248f).

Die algebraische Ordnung entspricht hier im Fall $g(A_2) < 1$ ¹⁾ wegen der Eigenwerteigenschaft von R aus S 2.10.a der Ordnungszahl μ , wenn auch für die Klasse $V^>$ das Ergebnis $\bar{v} \leq \mu$, $v = \min(\mu, m+1)$ - und nicht $\bar{v} = \mu$ - "sein wird". Die Charakterisierung der Konvergenzordnung für die Verfahrensklasse $V^>$ könnte man im Fall $g(A_2) < 1$ ähnlich zeigen wie z.B. bei Wanner/Hairer/Nørsett [45], S.486, womit allerdings weder die anschauliche Erklärung gewisser Effekte (Siehe Einleitung) noch die Charakterisierung der Ordnung für die Verfahren, welche nicht dem Dahlquistischen Wurzelkriterium genügen und "unter $g(A_2) = 1$ fallen", möglich wäre. Aus diesen Gründen soll hier der folgende Weg beschritten werden:

Zunächst ergibt sich eine Entwicklung der Form (3.3) aus der Darstellung (2.21) von R , wonach man die Konstanten c_2 und μ aus (3.3) mit c_1 und v aus (3.1) vergleichen kann. Die Charakterisie-

¹⁾ Aus (6.2) und S 6.5.c folgt : $S(W(0)) = \{1\} \cup S(A_2)$; $g(A_2) < 1$ und (2.4) ergeben das Erfülltsein des Dahlquistischen Kriteriums.

rungsaussage soll im Fall $q(A_2) = 1$ nur für $q = d_p + 1$ gezeigt werden, wozu noch einige Vorbetrachtungen notwendig sind. Sei also für das Folgende $q = d_p + 1$: Mit den Größen β, a, c, d aus K 2.5 sowie

$$(3.5) \quad \Delta := (a+d)^2 - 4a(d+c)$$

lauten die beiden Eigenwerte von Ψ und die beiden Fixpunkte von (2.12):

$$(3.6) \quad \begin{cases} \lambda_{\pm} := -\frac{1}{2\beta a} (a+d \pm \sqrt{\Delta}) & \text{(in } \mathbb{C} - N_{q_0}), \\ u_{\pm} := \frac{1}{2c} (a-d \pm \sqrt{\Delta}) & \text{(in } \mathbb{C} - N_{q_0} - N_{p_0} - N_c). \end{cases}$$

Weiter folgt mit $u := u_1$ aus (2.22) die Identität

$$(3.7) \quad u = \frac{1}{\beta R + 1} \quad (\text{in } K - F).$$

Man rechnet leicht nach, daß entsprechend (2.21) und (3.7) auch $\lambda_{\pm} = \tilde{g}(\cdot, (u_{\pm}, u_2), m)$ und $u_{\pm} = (\beta \lambda_{\pm} + 1)^{-1}$ gelten.

Bekanntlich lassen sich die Zweige der durch die Gleichung (2.13) definierten algebraischen Funktion λ in jedem einfach zusammenhängenden Gebiet in K , welches keine Verzweigungspunkte enthält, definieren, und sie sind dort holomorph bis auf ihre Pole. Im jetzt betrachteten Fall $q = d_p + 1$ seien die beiden Zweige von λ mit λ_0 und λ_1 bezeichnet, zum betrachteten Gebiet folgt sofort näheres. Zunächst jedoch noch die Bemerkung, daß im Fall einer algebraischen Gleichung -wie hier- zweiter Ordnung mit (2.14) die Nullstellen der Diskriminante Δ genau die endlichen Verzweigungspunkte sind. Gemäß S 6.4 gilt $\#N_{\Delta} < \infty$, d.h. zu jeder Parameterkonfiguration $J^{[0]} \in V^>$ mit $q = d_p + 1$ existiert eine (einfach zusammenhängende) Umgebung $W := W(0) \subset \mathbb{C}$ mit $N_{\Delta} \cap W \subset \{0\}$ (Die Schreibweise $W := W(z_0) \subset \mathbb{C}$ besage $z_0 \in \text{int}(W) \subset \mathbb{C}$). In dieser Umgebung sei vereinbarungsgemäß λ_0 der Hauptzweig, d.h. derjenige Zweig der Funktion λ mit $\lambda_0(z) \approx z^2$, $z \in W$. Nach dem oben Gesagten bleibt zu begründen, warum diese Definition in W auch dann möglich bleibt, wenn $z = 0$ Verzweigungspunkt ist, d.h. wenn $0 \in N_{\Delta}$ ²⁾ gilt. Der Grund ist der

1) $\sqrt{z} := +\sqrt{|z|} \cdot e^{i \arg(z)/2}$, $\sqrt{z} := -\sqrt{z}$, $z \in \mathbb{C}$.

2) Dies ist genau für $\beta = -1$ der Fall.

folgende: Man kann sich Überlegen, daß im Fall $\Delta(0)=0$ zudem stets

$$\Delta'(0) = 0, \quad \Delta''(0) = 2 \frac{(m+1)^2}{m^2} \neq 0$$

gelten, womit die Holomorphie der beiden Wurzeln und die Definierbarkeit der Zweige in $z = 0$ "ausnahmsweise" erhalten bleiben.

Die Zweige der (2.12) zugeordneten algebraischen Fixpunktgleichung $u = \varphi(u, z)$ seien mit $\underline{u}_0, \underline{u}_1$ bezeichnet und vermöge

$$\underline{u}_q := (\beta \cdot \lambda + 1)^{-1}$$

in $\mathbb{C}$ "numeriert". Man beachte $\{\lambda_0, \lambda_1\} = \{\lambda_+, \lambda_-\} = \{\lambda_0, \lambda_1\}$ sowie $\underline{u}_q = \tilde{g}(\cdot, (\underline{u}_0, \underline{u}_1), m)$.

Für die Charakterisierung der Konvergenzordnung im Fall $(q = d_0 + 1)$, $g(A_2) = 1$ kommt es wesentlich darauf an, die Taylor- bzw. Laurent-entwicklungen der Größen $\underline{u}_q, \underline{u}_q$ um $z = 0$ zu kennen, daher das folgende Lemma:

3.2 Lemma

Man betrachte zu $\beta \in \mathbb{C}$ mit $q = d_0 + 1$ die sukzessive Substitution (2.12) aus K 2.5. Es gelten die folgenden Aussagen:

- Für alle $z \in K - F$ besitzt (2.12) genau einen attraktiven Fixpunkt und konvergiert gegen diesen.
- Es existiert eine Umgebung $W := W(0) \subset \mathbb{C}$ mit $W \cap (K - F) \supset W - A' - \{0\}$ und es gelten die folgenden Entwicklungen:

$$(3.8) \quad \begin{cases} \underline{u}_0(z) = 1 + z + O(z^2), \\ \underline{u}_1(z) = -\frac{1}{\beta} \left(1 - \frac{z}{m}\right) + O(z^2), \quad z \rightarrow 0, \quad z \in W, \end{cases}$$

$$(3.9) \quad \begin{cases} \underline{u}_0(z) = \frac{1}{\beta(1+z)+1} + O(z^{2\text{sign}|\beta+1|}), \\ \underline{u}_1(z) = \frac{m}{z} + O(1), \\ \varphi'(\underline{u}_0(z), z) = -\frac{1}{\beta} \left(1 - \frac{m+1}{m} z\right) + O(z^2), \\ \varphi'(\underline{u}_1(z), z) = -\beta \left(1 + \frac{m+1}{m} z\right) + O(z^2), \quad z \rightarrow 0, \\ z \in W - \{0\}. \end{cases}$$

Korollar

Es gelten die Identitäten

$$(3.10) \quad R = \begin{cases} {}_0\lambda = \tilde{Q}(\cdot, ({}_0u, u_2), m) \text{ in } \{z \in W-A' - \{0\} \mid |\varphi'({}_0u(z), z)| < 1\}, \\ {}_1\lambda = \tilde{Q}(\cdot, ({}_1u, u_2), m) \text{ in } \{z \in W-A' - \{0\} \mid |\varphi'({}_1u(z), z)| < 1\}. \end{cases}$$

Beweis des Korollars

L 3.2.a besagt: Für die eine implizite Kollokationsstelle u , die Grenzwert von (2.12) ist, gilt $u = {}_0u$, sofern $|\varphi'({}_0u, \cdot)| < 1$ ist oder $u = {}_1u$, sofern $|\varphi'({}_1u, \cdot)| < 1$ gilt. Die Behauptung folgt aus (2.21), also aus der Identität $R = \tilde{Q}(\cdot, (u, u_2), m)$ in $K-F$.

Bemerkung

Aufgrund der Bedeutung der impliziten Kollokationsstelle (Siehe Bemerkung b zu T 2.8) ist die Tatsache, daß ${}_1u$ stets und ${}_0u$ für $\beta = -1$ einen Pol in $z = 0$ besitzen, eine "böse Überraschung". Aber genau darin liegt, wie im Anschluß an den angestrebten Hauptsatz genau ausgeführt wird, eine anschauliche Erklärung für das in $V^>$ auftretende Divergenz- bzw. "Ordnungsdefekt"-Phänomen.

Beweis des Lemmas

Der Einfachheit halber werde im folgenden (wieder) das Argument z weggelassen.

a) 1. Fall: $c = 0$

Wegen $z \in K-F$ ist $z \in K-A'$. Mit (3.6) folgt nach Definition von A' $|a| \neq |d|$. Ebenso gilt $z \in K-C$, was mit der Charakterisierung der Menge C im Fall $q = d_0 + 1$ (S 6.4) $|a| < |d|$ impliziert. (2.12) für $c = 0$ ergibt die Behauptung.

2. Fall: $c \neq 0$

Hier gilt für die Ableitungen der beiden Fixpunkte von (2.12)

$$(*) \quad \varphi'(u_{\pm}, \cdot) = \frac{4a(d+c)}{(a+d \pm \sqrt{\Delta})^2} \quad \text{und} \quad \varphi'(u_+, \cdot) \cdot \varphi'(u_-, \cdot) = 1.$$

Wiederum wegen $z \in K-A'$ folgt mit (3.6) $|a+d+\sqrt{\Delta}| \neq |a+d-\sqrt{\Delta}|$, d.h. mit (*): Genau einer der beiden Fixpunkte ist attraktiv.

Aus T 2.8 und (3.7) folgt: $u_k \xrightarrow[k \rightarrow \infty]{} u$ mit $u = (\beta R + 1)^{-1}$.

$R = \lambda_0$, (3.6), $\{\lambda_+, \lambda_-\} = \{\lambda_0, \lambda_1\}$ sowie $|\lambda_0| > |\lambda_1|$ zeigen mittels (*) $|\varphi'(u, \cdot)| < 1$, also die Attraktivität des Grenzwertes der

Folge $(u_k)_{k \in \mathbb{N}_0}$.

- b) Die Existenz einer Umgebung $W := W(0) \subset \mathbb{C}$ mit $W \cap (K-F) \supset W-A' - \{0\}$ folgt unmittelbar aus D 2.7 sowie der Endlichkeit der Mengen N_{q_0} , N_{p_0} , B und C (Vgl. S 6.4).

Die Taylor- bzw. Laurententwicklungen um $z = 0$ sind zunächst nur für die Größen ${}_0\lambda$ und $\varphi'({}_0u, \cdot)$ durchzuführen, wobei man die Fälle $\beta \neq -1$, $\beta = -1$ zu unterscheiden hat. Bzgl. ${}_0\lambda$ ist dazu in (3.6) dasjenige Vorzeichen von $\sqrt{\Delta}$ zu wählen, was zu ${}_0\lambda(z) \approx e^z$, $z \in W$, führt. Das gleiche Vorzeichen ist "für" $\varphi'({}_0u, \cdot)$ in (*) zu wählen. Die weiteren Entwicklungen ergeben sich auf elegantere Art: Nach der Vietaschen Formel für das Absolutglied einer quadratischen Gleichung folgt mit S 6.3.b(ii)

$${}_1\lambda = {}_0\lambda^{-1} \det \Psi = -(\beta_0 \lambda)^{-1} \frac{p_0^c}{q_0} ;$$

desweiteren gilt wegen (*) einerseits und des Zusammenhangs zwischen den Zweigen von (2.12) und (2.13) (Vgl. (3.7)) andererseits

$$\varphi'({}_1u, \cdot) = \varphi'({}_0u, \cdot)^{-1}, \quad {}_1u = (\beta_1 \lambda + 1)^{-1}.$$

Vor der Formulierung des Hauptsatzes muß noch die Identität aus T 1.2.a betrachtet mit $W := U^{[0]} \in V^{\leq}$ auf die Klasse $V^>$ bzw. auf die Funktion R aus (2.21) verallgemeinert werden:

3.3 Lemma

- a) Sei $U^{[0]} \in V^>$ und dazu gemäß (1.5), (2.3) und T 2.8

$$(3.11) \quad \begin{cases} \tilde{N}(t) &:= t^d \prod_{i=1}^p (t-t_i)^{d_i} , \\ N(t, z) &:= v(t)^T x(z) , \quad t \in \mathbb{R}, \quad z \in K-F . \end{cases}$$

Dann gilt in Verallgemeinerung von Formel (5a) aus Nörsett/Wanner [32] die Identität

$$(3.12) \quad e^z - R(z) = \frac{z^{m+1}}{m!q(z)} \int_0^1 e^{z(1-\tau)} N(\tau, z) d\tau, \quad z \in K-F.$$

- b) Sei $U^{[0]} \in V^>$ mit $q(A_2) < 1$ oder $q(A_2) = 1$ und $q = d_p + 1$.

Dann existiert eine Umgebung $W := W(0) \subset \mathbb{C}$ mit $W \cap \mathbb{R}_+ \cap F = \emptyset$ (sogar mit $W \cap F = \emptyset$ im Fall $q(A_2) < 1$).

Bemerkung

Die Identität (3.12) erlaubt es insbesondere, höhere Ordnungen $\mu > m$ -ähnlich wie im Fall $q \leq d_p$ - mittels (integraler) Orthogonalitäten zu charakterisieren. Hierzu benötigt man nurmehr die Taylorentwicklung von x_r um $z = 0$, welche man mittels (2.18) und (3.12) selbst leicht erhält. Dies soll jedoch hier nicht weiter ausgeführt werden.

Beweis

a) Völlig analog zum Beweis von Formel (5a) in Nørsett/Wanner [32], wobei zu festem $z \in \mathbb{K} - F$ als Kollokationsstellen die m Komponenten von $U(z)$ aus (2.17) genommen werden. Hierbei gilt das im Beweis zu S 2.4.a, Formel (2.9), gesagte entsprechend und es ist $Q(z) \neq 0$, $z \in \mathbb{K} - F$ -siehe (2.19)- zu beachten.

b) 1.Fall: $Q(A_2) = 1$, $q = d_p + 1$

Es ist $|\beta| = 1$ (S 6.5.c, Beispiel a zu S 6.5). L 3.2.b, (3.8) implizieren: Es existiert eine Umgebung $W := W(0) \subset \mathbb{C}$ mit $W \cap (\mathbb{K} - F) \supset W - A' - \{0\}$ und $|\lambda_0(z)| > |\lambda_1(z)|$, $z \in W \cap \mathbb{R}_+$; folglich gilt $W \cap A' \cap \mathbb{R}_+ = \emptyset$, was $W \cap (\mathbb{K} - F) \cap \mathbb{R}_+ = W \cap \mathbb{R}_+$ impliziert.

2.Fall: $Q(A_2) < 1$

Aus den Formeln (6.2) und (2.2) sowie S 6.5.c folgt $S(\Psi(0)) = \{1\} \cup S(A_2)$. Demnach gilt für die Eigenwerte der Matrix $\Psi(0)$:

$$\lambda_0(0) = 1, |\lambda_j(0)| < 1, j = 1(1)q - d_p.$$

Es ist weiter

$$s_k(0) = 1 \cdot (\lambda_0(0))^k + \sum_{j=1}^{q-d_p} 0 \cdot (\lambda_j(0))^k, k \in \mathbb{N}_0.$$

Für den zu $\lambda_0(z)$ gehörenden Eigenvektor $y_0(z)$ gilt bei der in $z=0$ möglichen Normierung $e_0^T y_0(z) = 1$

$$y_0(0) = (1, ((H - H^*)^{-1} h^*)^T)^T.$$

Die hierbei auftretende Matrix ist wegen $Q(H^{-1} h^*) = Q(A_2) < 1$ invertierbar (Vgl. S 6.5.c).

Aufgrund der Stetigkeit der Abbildung $z \mapsto (|\lambda_0(z)|, \dots, |\lambda_{q-d_p}(z)|)$, der stetigen z -Abhängigkeit des Beteiligungskoeffizienten von λ_0 in der Vektoriteration (I) aus 2-2 bei Nichtverzweigung zu einem anderen Eigenwert sowie der von y_0 -wie oben normiert- folgt unter Beachtung von D 2.7 und der Bemerkung a zu S 6.4

die Existenz einer Umgebung $\tilde{W} := \tilde{W}(0) \subset \mathbb{C}$ mit $\tilde{W} \cap (A' \cup B \cup C) = \emptyset$. Die Überlegungen sollen im einzelnen nicht ausgeführt werden, nur eine Bemerkung zur Stetigkeit des Beteiligungskoeffizienten: Er ist das Skalarprodukt der zugeordneten Zeile der Inversen einer (ggf. konfluenten) Vandermondematrix¹⁾ mit dem Startvektor von (I). Unstetig sind nur die verzweigenden Eigenwerten zugeordneten Zeilen dieser Inversen. Der Rest ist mit D 2.7 trivial.

3.4 Theorem

Sei $U^{[0]} \in V^>$ mit $g(A_2) < 1$ oder

$g(A_2) = 1$ und $q = d_p + 1$ oder

$g(A_2) > 1$ und $0 \notin \overline{B - \{0\}}$ sowie $t_i \in (0, 1)$, $i = 1(1)p-1$.

Dann wird die Ordnung ν von $U^{[0]}$ durch die Ordnung μ des zugehörigen R mit einer einzigen Ausnahme in der folgenden Weise charakterisiert:

$$\nu = \min(\mu, m+1)$$

Im Fall $q = d_p + 1$, $\beta = -1$, $\int_0^1 \tilde{N}(\tau) d\tau = 0$ ist $\nu = m, \mu > m+1$.

Bemerkungen

- Für den Fall $q = d_p + 1$ ist die obige Aussage (ohne Beweis) bereits in dem Tagungsbericht Fuchs [8] (Charakterisierungssatz) veröffentlicht.
- Für die Klasse der linearen MSV mit höheren Ableitungen gilt (bei Erfülltsein des Dahlquistischen Wurzelkriteriums) $\bar{\nu} = \mu$. Bzgl. $V \leq$ bzw. für die entsprechende Teilklasse der IRK-Methoden gilt wie oben $\nu = \min(\mu, m+1)$ und auch $\bar{\nu} = \mu$ (Burrage [1], (9), Nørsett/Wanner [32], (5b), Mülthei [29], Satz 4 und [28], Satz 3 und Korollar 3). Für die Klasse $V^>$ kann sogar $\bar{\nu} < \mu$ sein, was hier nicht bewiesen werden soll.
- Wie bereits gesagt, erlaubt die hiesige Vorgehensweise auch die Charakterisierung der Konvergenzordnung bei solchen Repräsentanten von $V^>$, die nicht dem Dahlquistischen Wurzelkriterium genügen. Exemplarisch wird dies nur für die Fälle mit $q = d_p + 1$ durchgeführt. Für die Beispiele mit $q > d_p + 1$ hätte man die Laurententwicklungen der verschiedenen Zweige der (symmetrischen Funktionen der) impliziten Kollokationsstellen entsprechend

¹⁾ gebildet aus den Eigenwerten von $V(z)$ (Vgl. Gautschi [47], (3.2), (4.6)).

(3.9) zu bestimmen (Es ist zu vermuten, daß es Verfahren mit $\nu = m - (q - d_p)$ gibt).

- d) Eine Erklärung für den angesprochenen einzigen Ausnahmefall folgt in B 3.5.b.

Beweis

Die Charakterisierungsaussage soll also durch Vergleich der Ordnungen ν von V und μ von R erhalten werden.

1. Fall: $Q(A_2) > 1$

Aus Mülthei [26], Satz 5¹⁾ folgt $\nu = -1$. Wegen $0 \in \overline{B - \{0\}}$ existiert eine Umgebung $W := W(0) \subset \mathbb{C}$ mit $W \cap (K - B) \supset W - \{0\}$; die Formeln (2.10) und (6.2) sowie S 6.5.c zeigen: $\lim_{\substack{z \rightarrow 0 \\ z \neq 0}} |R(z)| = |\lambda_0(0)| = Q(A_2) > 1$.
D 3.1.c $\Rightarrow \mu = -1$.

2. Fall: $Q(A_2) = 1$

(Es wird $q = d_p + 1$ betrachtet, und es ist damit $|\beta| = 1$)

Die Konvergenzordnungen $m-1, m, m+1$ sind durch gewisse Bedingungen an die Matrizen $\tilde{A}$ (Vgl. Mülthei [26], Satz 4, [27], Satz 2 und [28], Satz 4) und $\tilde{A}_z$ ([28], Satz 4) charakterisiert. Mit den hier eingeführten Bezeichnungen aus (2.11), (3.11) und (6.3) gelten

$$\tilde{A} = \begin{pmatrix} 1 & -\frac{1}{\varepsilon \gamma_0} \int_0^1 \tilde{N}(\tau) d\tau \\ 0 & -1/\beta \end{pmatrix}, \tilde{A}_z = \begin{pmatrix} \frac{-1}{\varepsilon \gamma_0 d_p!} \int_0^1 \tilde{N}(\tau) d\tau & \frac{1}{m!} \int_0^1 \tau \tilde{N}(\tau) d\tau \\ -1/\beta & -1 \end{pmatrix}, \tilde{A}_z = \begin{pmatrix} \frac{d_p!}{m!} \gamma_0^* & \end{pmatrix}.$$

Es folgen die Bedingungen:

$$(*) \quad \begin{cases} \nu = m-1 \iff \beta = -1, \int_0^1 \tilde{N}(\tau) d\tau \neq 0, \\ \nu = m \iff \beta = -1, \int_0^1 \tilde{N}(\tau) d\tau = 0 \text{ oder } \beta = 1, \det \tilde{A}_z \neq 0, \\ \nu = m+1 \iff \beta = 1, \det \tilde{A}_z = 0. \end{cases}$$

Wegen L 3.3.b, (3.9) und $|\beta| = 1$ existiert eine Umgebung $W := W(0) \subset \mathbb{C}$ derart, daß die Identität (3.10) in $W \cap \mathbb{R}_+$ stets "mit u " gilt. Mittels L 3.3.b und a ergibt sich für $z \in W \cap \mathbb{R}_+$

¹⁾ Die Matrix A_2 wurde in dieser Arbeit noch mit A bezeichnet.

$$e^z - R(z) = \frac{z^{m+1}}{m! q(z)^T (1, {}_0u(z))^T} \int_0^1 e^{z(1-\tau)} \tilde{N}(\tau) (\tau - {}_0u(z)) d\tau.$$

Aus der Definition von q in (2.2) sowie der Entwicklung von ${}_0u$ in (3.9) folgt für den hiesigen Fall $g(A_2) = 1$, $q = d_p + 1$:

$$\begin{aligned} e^z - R(z) &= \frac{z^m}{(m-1)!(m+1)} \int_0^1 \tilde{N}(\tau) d\tau + O(z^{m+1}), z \rightarrow 0, z \in \mathbb{R}_+, \beta = -1, \\ (**) \quad e^z - R(z) &= \frac{z^{m+1}}{m!} \int_0^1 \tilde{N}(\tau) (\tau - \frac{1}{2}) d\tau + O(z^{m+2}), z \rightarrow 0, z \in \mathbb{R}_+, \beta = 1. \end{aligned}$$

Hier sei schon darauf hingewiesen, daß für $\beta = -1$ der Pol von ${}_0u$ mit ${}_0u(z) = -1/z + O(1)$ im "Ordnungspunkt" $z = 0$ zum Ordnungsdefekt führt (Siehe B 3.5.b). Im Fall $\beta = 1$ gilt wegen $\beta = \varepsilon \gamma_0 / \gamma_0^*$

$$\det \tilde{A}_z = \frac{1}{m!} \int_0^1 \tilde{N}(\tau) (2\tau - 1) d\tau.$$

Die Aussage folgt mittels D 3.1.c durch Vergleich von (*) mit (**), wobei im angesprochenen Ausnahmefall $\mu > m+1$ gilt (Man beachte hierbei $\exp(z(1-\tau)) \tilde{N}(\tau) (\tau - {}_0u(z)) = (z^{-1} + 1/2 + O_\varepsilon(z)) \tilde{N}(\tau)$, $z \rightarrow 0, z \in \mathbb{R}_+$).

3.Fall: $g(A_2) < 1$

Aus den Aussagen von Mülthei folgt $\nu \geq m$. Weiter ist für die genaue Konvergenzordnung (wieder) die Matrix $\tilde{A}_z$ ([28], S.457) verantwortlich. Zunächst sollen die folgenden 3 Zwischenbehauptungen bewiesen werden.

(i) $\text{rang } \tilde{A}_z \geq q - d_p,$

(ii) $\nu = m+1 \Leftrightarrow \det \tilde{A}_z = 0,$

(iii) $\det \tilde{A}_z = \frac{1}{m!} \det(I_r - H^{-1} H^*) \int_0^1 (v_0(\tau) + v_r(\tau))^T (I_r - H^{-1} H^*)^{-1} H^{-1} h^* d\tau.$

Zu (i): Aus der Definition von $\tilde{A}_z$ in Mülthei [28], S.457, ersieht man, daß $\tilde{A}_z - A_2 - I_r$ als Teilmatrix enthält. Aus $g(A_2) < 1$ folgt (i).

Zu (ii): Für 1 aus [28], Satz 4 gilt wegen $g(A_2) < 1$ $l = 1$, laut (4.4) in [28] folgt die Äquivalenz " $\nu = m+1 \Leftrightarrow \text{rang } \tilde{A}_z = q - d_p$ ". Die Behauptung unter (ii) ergibt sich nun unmittelbar aus (i).

Zu (iii): Zunächst die Bemerkung, daß die auftretende Matrix $I_r - H^{-1} H^*$ wegen $g(H^{-1} H^*) = g(A_2) < 1$ regulär ist und ebenso $\det(I_r - H^{-1} H^*) \neq 0$ gilt (Vgl. S 6.5.c). Mit den Bezeichnungen z_{1p} , α_p , z_2 (und A_2) aus [28] sowie mit S 6.5.b und c folgt die Identität

$$\tilde{A}_Z^I := \begin{pmatrix} z_{1p} & \alpha_p \\ z_2 & A_2 - I_r \end{pmatrix} = \begin{pmatrix} \frac{1}{m!} \int_0^1 v_0(\tau) d\tau & \int_0^1 v_r(\tau)^T H^{-1} d\tau \\ \frac{1}{m!} h^* & H^* H^{-1} - I_r \end{pmatrix}.$$

Mit $\det \tilde{A}_Z = (-1)^{q-d_p} \det \tilde{A}_Z^I$ und L 6.2 ist (iii) bewiesen.

L 3.3 besagt für den Fall $q(A_2) < 1$ die Entwicklung

$$e^{z-R}(z) = \frac{z^{m+1}}{m!} \int_0^1 v(\tau)^T x(0) d\tau + O(z^{m+2}), \quad z \rightarrow 0,$$

was mit D 3.1.c $\underline{\mu} \geq m$ ergibt ; (2.18) für $z=0$, $x_0(0)=1$, (ii) und (iii) implizieren die zum Abschluß des Beweises noch fehlende Äquivalenz: $\nu = m+1 \Leftrightarrow \underline{\mu} \geq m+1$. ●

Korollar

a) In den Fällen mit $q(A_2) < 1$ oder $q(A_2) = 1$, $q=d_p+1$, $\beta=+1$ ¹⁾ kann man die Integralbedingung für die maximale Konvergenzordnung $\nu = m+1$ aufgrund der Charakterisierungsaussage wie folgt schreiben:

$$(3.13) \quad \int_0^1 \prod_{j=1}^{q-d_p} (\tau - u_j(0)) \tau^{d_p} \prod_{i=1}^p (\tau - t_i)^{d_i} d\tau = 0$$

(Hierzu sei im Fall $q=d_p+1$, $\beta=+1$, $u_1(0) := u_1(0+) := \lim_{\substack{z \rightarrow 0 \\ z \in \mathbb{R}_+}} u_1(z)$).
Dabei ist die obige Bedingung vermöge

$$(3.14) \quad x_0(0) = 1, \quad x_r(0) = (I_r - H^{-1}H^*)^{-1}H^{-1}h^*$$

durch die Verfahrensparameter charakterisiert (Man beachte $x(0) = (\sigma_\alpha(u_1(0), \dots, u_{q-d_p}(0))), \alpha=0(1)q-d_p$) und formelmäßig natürlich in der Aussage von Mülthei [28], Satz 4, enthalten.

Interessant ist hier jedoch die formale Analogie von (3.13) zu der entsprechenden Bedingung (1.8) für die Klasse $V^{\leq}$, was nach dem Über die impliziten Kollokationsstellen Gesagten nun nicht mehr verwundert. Bemerkung b zu T 2.8 besagt nämlich gerade, daß das Verfahren nach "genügend langer Rechnung" für kleine Schrittweiten (approximativ) mit den Kollokationsstellen $U(0) = (u_1(0), \dots, u_{q-d_p}(0), u_2)$ "arbeitet". Interessant ist zudem, daß

¹⁾ Dahlquistisches Wurzelkriterium (für $S(\Psi(0))$) ist erfüllt.

das für die Testgleichung eingeführte Konzept der freien und impliziten Kollokationsstellen seine Gültigkeit für allgemeine rechte Seiten f in (1.1) offensichtlich zumindest im Bereich kleiner Schrittweiten nicht verliert (Auf den Fall $\beta = -1$ sowie die in der Einleitung versprochene anschauliche Erklärung für einen gewissen Effekt wird -wie gesagt- sogleich näher in § 3.5 eingegangen).

- b) Bei Divergenz im Fall $q = d_p + 1$ ($|\beta| < 1$) existiert wegen L 3.2 und (3.10) eine Umgebung $W := W(0) \subset \mathbb{C}$ mit

$$u(z) = {}_1u(z) = \frac{m}{z} + O(1),$$

$$R(z) = {}_1\lambda(z) = \tilde{g}(z, ({}_1u(z), u_2), m) = -\frac{1}{\beta} + O(z), z \rightarrow 0, z \in W - \{0\}.$$

In diesem Fall zerstört also die eine implizite Kollokationsstelle die im "Ansatz" $\tilde{g}$ steckende Ordnung m völlig.

3.5 Bemerkungen

Zu $U^{[0]} \in V$ sei $(w_1, \dots, w_m) := U^{[0]}$.

- a) Bekanntlich gilt für $U = U^{[0]} \in \underline{V}$ (Vgl. (1.5), (1.8), Bemerkung b zu T 3.4, T 1.2.a):

$$v \geq m \quad \text{und} \quad "v = m+1 \iff \int_0^1 \prod_{i=1}^m (\tau - w_i) d\tau = 0"$$

- b) Für $U^{[0]} \in \underline{V}^>$ können sowohl der Fall $v = m+1$ mit $\int_0^1 \prod_{i=1}^m (\tau - w_i) d\tau \neq 0$ als auch $v = m-1$ auftreten, was von Mülthel für $p=1$ in [24], Satz 5 und Nachtrag, bzw. [25], Satz 2, gezeigt wurde.

Dieses Phänomen sowie die in T 3.4 angesprochene Ausnahme, wo $v = m$, jedoch $\mu \geq m+1$ gilt, lassen sich mit der aufgezeigten Theorie der impliziten Kollokationsstellen plausibel machen. Bei den beiden Ausnahmen handelt es sich durchaus um zwei verschiedene Effekte, nämlich einmal um höhere Konvergenzordnung trotz der Eigenschaft des Parametersatzes, daß $\int_0^1 \prod_{i=1}^m (\tau - w_i) d\tau \neq 0$ gilt, zum anderen um "Konvergenzordnungsdefekt".

Definiert man -wie oben- auch für $q = d_p + 1, \beta = -1, u_1(0) := u_1(0+) (= \infty)$, so zeigen (3.13) und T 3.4, daß in allen behandelten Fällen mit $Q(A_2) \leq 1$ bzgl. der Konvergenzordnung in $\underline{V}^>$ an die Stelle von $U \in \mathbb{R}^m$ der Vektor

$$U(0) = (u_1(0), \dots, u_{q-d_p}(0), u_2) \in \mathbb{R}^m$$

tritt. Man muß sich nur noch daran erinnern, daß die Nörsett/Wannersche Funktion $\tilde{g}$ die Eigenschaft besitzt, daß Kollokationsstellen im Unendlichen genau wie nicht vorhandene zu betrachten sind ¹⁾, mit anderen Worten: Jedem $U^{[0]} \in V^>$ ist in natürlicher Weise ein Vektor $\underline{U} \in \mathbb{R}^{m'}$ zugeordnet, den man erhält, indem man aus $U(0)$ den Wert ∞ , soweit vorhanden, herausstreicht. In allen (behandelten) Fällen mit $g(A_2) \leq 1$ gilt die Regel für $V^<$ unter a entsprechend auch hier:

$$v \geq m' \quad \text{und} \quad " \quad v = m' + 1 \iff \int_0^1 \prod_{i=1}^{m'} (\tau - \omega_i') d\tau = 0 \quad "$$

(Hierbei ist $(\omega_1', \dots, \omega_{m'}') := U'$). Insbesondere ist hiermit die Ausnahme in T 3.4 geklärt, stets gilt $v = \min(\mu, m' + 1)$. Diese Ausnahme sowie der bereits in der Einleitung angesprochene Fall $p = 1$ seien noch einmal genauer skizziert:

$$q = d_p + 1, \beta = -1$$

$$U^{[0]} = (0, \underbrace{0, \dots, 0}_{d_p}, \underbrace{1, \dots, 1}_{d_p}) \in \mathbb{R}^m; \quad U' = (\underbrace{0, \dots, 0}_{d_p}, \underbrace{1, \dots, 1}_{d_p}) \in \mathbb{R}^{m'}$$

mit $m' = m - 1$ ($u_1(0) = \infty$). Die obige Regel besagt (mit (3.11)):

$$v \geq m - 1 \quad \text{und} \quad " \quad v = m \iff \int_0^1 \tilde{N}(\tau) d\tau = 0 \quad "$$

$$q = d_p + 1, p = 1 \text{ (Vgl. [24], [25])}$$

$$U^{[0]} = (\underbrace{0, \dots, 0}_{d_1+1}, \underbrace{1, \dots, 1}_{d_1}) \in \mathbb{R}^m, \beta = (-1)^{d_1+1}$$

(i) d_1 gerade

Aus [25], Satz 2 folgt $v = m - 1$; Erklärung hier:

$$\beta = -1, u_1(0) = \infty \Rightarrow U' = (0, \dots, 0, 1, \dots, 1) \in \mathbb{R}^{m'},$$

$$m' = m - 1,$$

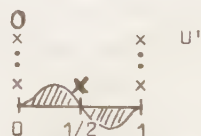
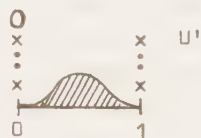
$$\int_0^1 \tau^{d_1} (\tau - 1)^{d_1} d\tau \neq 0.$$

(ii) d_1 ungerade

Aus [24], Satz 5 folgt $v = m + 1$; Erklärung hier:

$$\beta = +1, u_1(0) = 1/2 \Rightarrow U' = (0, \dots, 0, 1/2, 1, \dots, 1) \in \mathbb{R}^{m'},$$

$$\int_0^1 \tau^{d_1} (\tau - 1/2) (\tau - 1)^{d_1} d\tau = 0.$$



¹⁾ Eine entsprechende Verallgemeinerung von L 1.3 ist möglich genau, falls $u(z) = cz^{-n} + O(z^{1-n}), z \rightarrow 0, (z \in \mathbb{R}_+, c \neq 0, c \neq m \text{ oder } n > 1)$, gilt.

Die eingezeichneten, für den "Ordnungspunkt" $z = 0$ relevanten Parameterkonfigurationen U' liefern die versprochene anschauliche Erklärung für die beiden möglichen Konvergenzordnungen $v = m-1$ und $v = m+1$ im Fall $p = 1$.

An dieser Stelle sei die Bemerkung gestattet, daß die in der vorliegenden Arbeit bislang erläuterte Theorie ausgehend vom Verhalten der einen freien Kollokationsstelle in den beiden obigen Fällen mit $p = 1$ entwickelt wurde.

3.6 Beispiele

Die folgenden Beispiele sollen demonstrieren, wie man mit den hier eingeführten Größen vermöge (3.13), (3.14) Parameterkonfigurationen mit $v = m+1$ finden kann. Die beiden eben erwähnten Formeln, S 6.5.c sowie die Beispiele a und b zu S 6.5 zeigen für die Fälle

$$q = d_p + 1$$

$$(3.15) \quad \left\{ \begin{array}{l} g(A_2) = |\beta|^{-1} \quad ; \\ v = m+1 \iff \int_0^1 (\tau - u_1(0)) \tilde{N}(\tau) d\tau = 0, \text{ mit} \\ u_1(0) = (\beta + 1)^{-1} \end{array} \right.$$

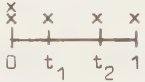
$$q = d_p + 2$$

$$(3.16) \quad \left\{ \begin{array}{l} g(A_2) = \begin{cases} |\beta|^{-1} & , \Gamma \in [-4, 0] , \\ \left| \frac{1}{\beta} \cdot \frac{1 + \sqrt{1+4/\Gamma}}{1 - \sqrt{1+4/\Gamma}} \right| & , \Gamma \notin [-4, 0] , \text{ mit} \end{cases} \\ \Gamma := 2d_p + \sum_{i=1}^{p-1} \frac{d_i}{t_i(1-t_i)} \quad ; \\ v = m+1 \iff \int_0^1 (\tau - u_1(0))(\tau - u_2(0)) \tilde{N}(\tau) d\tau = 0, \text{ mit} \\ (u_1(0) + u_2(0), u_1(0)u_2(0)) = \delta^{-1}(\delta - \beta_+ \beta_-, \beta_+) , \\ \beta_{\pm} := \beta_{\pm} - 1 \quad , \quad \delta := \beta_-^2 - \Gamma \beta \end{array} \right.$$

Hierbei ist für die Äquivalenzen mit $v = m+1$ und für die Darstellungen

von $u_1(0), u_2(0)$ $g(A_2) < 1$ unterstellt und für $q = d_p + 1$ noch $\beta = +1$ zugelassen. Nun also zu zwei konkreten Fällen:

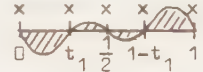
a) $q = d_p + 1$ (Vgl. Mülthei [28], Beispiel 7, S. 461 f):

$$p=3, q=2, d_3=1; d_1=d_2=1, t_1, t_2 > 0, t_1 \neq t_2, t_1, t_2 \neq 1$$


Genau für die Fälle $t_1 = 1/2, t_2 \in (1/2, \infty) - \{1\}$ oder $t_2 = 1/2, t_1 \in (1/2, \infty) - \{1\}$ oder $t_1 + t_2 = 1, t_2 \in (0, 1) - \{1/2\}$ ist die Bedingung für $v = m+1 = 6$ erfüllt (und mit $\det \tilde{A}_2 = 0$ äquivalent, vgl. [28]).

In der hiesigen Sprache der impliziten Kollokationsstellen findet das folgende "Wechselspiel" statt: $t_1 = 1/2 \Rightarrow u_1(0) = 1 - t_2$, $t_2 = 1/2 \Rightarrow u_1(0) = 1 - t_1$; $t_1 + t_2 = 1 \Rightarrow u_1(0) = 1/2$. In allen Fällen mit $v = 6$ gilt folglich (Vgl. 8.3.5.b) :

$$U' = (0, t_1, 1/2, 1 - t_1, 1)$$

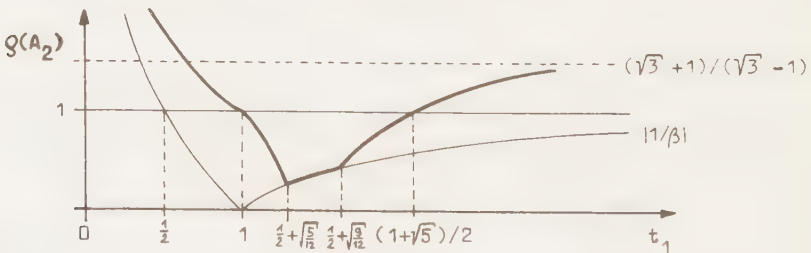


b) $q = d_p + 2$

$$p=2, q=3, d_2=d_1=1, t_1 > 0^1)$$



Für $g(A_2)$ ergibt sich etwa folgende Skizze:



Die Divergenz des Verfahrens für $t_1 \in (0, 1)$ läßt sich (auch) aus Satz 8.2 in Mülthei [26] schließen. Dabei zeigt das hiesige Beispiel, daß die Voraussetzung $t_1 \in (0, 1)$, $i=1(1)p-1$, notwendig für die dortige Divergenzaussage ist. Die Bedingung für maximale Konvergenzordnung $v = 6$ aus (3.16) ist, gemessen an den Definitionen der relevanten Größen β, Γ , überraschend einfach und lautet

1) Im Fall $t_1 = 1$ werde unter dem Verfahren der (konfluente) Repräsentant $U^{[0]} = (0, 0, 0, 1, 1)$ bzw. das Vf. mit $p=1, q=3, d_1=2$ verstanden.

$$(2t_1-1)(3t_1^2-3t_1-1) = 0, \quad t_1 \in \left(1, \frac{1+\sqrt{5}}{2}\right).$$

Folglich ist $t_1^S := 1/2 + \sqrt{7/12} \doteq 1.263762616$ (einziger) "Superkonvergenzparameter". Man beachte, daß in t_1^S $g(A_2) < 1$ gilt. Insgesamt folgt (Divergenz für $t_1 > (1+\sqrt{5})/2$ unterstellt):

ν	-1	4	5	6
t_1	$(0, 1) \cup \left(\frac{1+\sqrt{5}}{2}, \infty\right)$	1	$\left(1, \frac{1+\sqrt{5}}{2}\right) - \{t_1^S\}$	t_1^S

■ 3-1

Zum Abschluß dieses Kapitels noch eine Bemerkung, welche weitere Möglichkeit die Darstellung (2.21) bzw. die Identität (3.12) beinhalten: Zu $U^{[0]} \in V^>$ mit $g(A_2) < 1$ erhält man vermöge (2.18) und (3.12) sofort die "Entwicklung"

$$(3.17) \quad x_r(z) = (I_r e^z - H^{-1} H^*)^{-1} H^{-1} h^* + O(z^{m+1}), \quad z \rightarrow 0.$$

Dies liefert die Möglichkeit, höhere Ordnungen von R ($m+1 < \mu \leq 2m$) mittels integraler Orthogonalitäten zu charakterisieren. Vermöge der Darstellung (2.1) gelangt man schließlich zu Bedingungen, wann höhere Ordnungen in den Hauptknoten ($\bar{\nu} > m+1$) vorliegen.

Da (2.1) jedoch für die Anwendung von $U^{[0]}$ auf das spezielle Anfangswertproblem (T_μ) gilt, soll auf die Herleitung und Angabe jener Bedingungen verzichtet werden. Diese haben nämlich erst dann praktische Bedeutung, wenn sie auch für allgemeine rechte Seiten f in (1.1) ihre Gültigkeit behalten, was noch zu untersuchen ist.

■ 3

4 A_∞ -Stabilität

In diesem Kapitel sollen Parameterkonfigurationen, für die A_∞ -Stabilität vorliegt, durch einfach überprüfbare Kriterien charakterisiert werden. Bekanntlich sind A_∞ -stabile Verfahren in vielen Fällen "fast A-stabil" oder sogar A-stabil, worin unter anderem die Bedeutung obiger Eigenschaft liegt. Dies trifft zumindest dann zu, wenn (das Verfahren konvergent ist und) der Rand der Region absoluter Stabilität ∂R_A keine Doppelpunkte besitzt, was zwar häufig der Fall ist, jedoch so ohne weiteres nicht nachgewiesen werden kann (Vgl. z.B. Fuchs [7], Satz 1.4.11, für die Verfahren von Gear [9]).

4.1 Definition

Sei $U^{[0]} \in V$. $U^{[0]}$ heie A_∞ -stabil : $\Leftrightarrow \infty \in \text{int } \overline{R_A}^{1)$.

Bemerkungen

- a) D 4.1 besagt, da es im Fall der A_∞ -Stabilität eine Umgebung $X := X(\infty) \subset \mathbb{K}^{2)}$ gibt mit $X \setminus \{\infty\} \subset R_A$. Damit sind (gegenüber anderslautenden Definitionen) die Fälle ausgeschlossen, bei denen eine -existierende- Wachstumsfunktion R in $z = \infty$ betragsmäig den Wert 1 annimmt (Beispiel: $U = U^{[0]} = (0, 1)$, $R(\infty) = -1$ (Trapezregel)). Verfahren, die nach obigem Begriff A_∞ -stabil sind, besitzen die Eigenschaft, da bzgl. der Schrittweite h sehr steife Komponenten der numerischen Lösung eines AWP der Form (1.1)³⁾ noch "echt gedämpft" werden, womit die zweite Bedeutung der in D 4.1 definierten Eigenschaft genannt wäre.
- b) Zu einem wichtigen Unterschied in der Bedeutung der A_∞ -Stabilität zwischen $V^{\leq}$ und $V^>$, der entsprechend auch für die Eigenschaft A-Stabilität gilt, kann erst nach dem folgenden Satz etwas gesagt werden.

1) Unter dem Abschluß $\overline{M}$ einer Menge M werde stets der Abschluß in $\mathbb{K}$ verstanden.

2) Die Schreibweise besage $\infty \in \text{int } X \subset \mathbb{K}$.

3) Wenn von AWP der Form (1.1) im Zusammenhang mit steifen Systemen die Rede ist, so sei $f : [0, 1] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit $n \in \mathbb{N}$ zulässig.

4-1 Charakterisierung der A_∞ -Stabilität

Der Beweis des folgenden Satzes stellt die wesentliche Arbeit auf dem Weg zu den Aussagen dieses Kapitels dar.

4.2 Satz

Sei $U^{[0]} \in V^>$, dazu $U_P^{[0]} := (0, U^{[0]})^1$. Die $U_P^{[0]}$ zugeordneten Größen werden sämtlich mit einem (ggf. weiteren) Index P versehen. Die $U^{[0]}$ zugeordnete Matrix Ψ ist holomorph in $z = \infty$ mit

$$(4.1) \quad \Psi(\infty) = H_P^{-1} H_P^* \quad 2)$$

Korollar

Die Matrix $\Psi(\infty)$ ist regulär mit $\det \Psi(\infty) = \beta_P^{d_P - q - 1}$.

Beweis des Korollars

S 6.5 c, d mit $\beta_P = -\beta$, $q_P = q+1$

Beweis des Satzes

Im ersten Teil wird gezeigt, daß Ψ aus (2.2) in $z = \infty$ stetig ergänzbar ist (Die Holomorphie folgt dann aus der komponentenweisen Rationalität). Gemäß (2.2) bzw. (6.2) hat man nur Zeile Nummer 0 von Ψ zu untersuchen, wobei $\text{grad } p_\alpha = m - d_P$, $\text{grad } q_\alpha = m - q + \alpha$, $\alpha = 0(1)q - d_P$, vermerkt sei. Wegen $\text{grad } q_0 = m - q$ ist also zu zeigen, daß die Koeffizienten von z^{m-q+1} bis z^{m-d_P} in $p(z)^T$ und $q_T(z)^T H^{-1}(h^*, h^*)$ dieselben sind. Dies ist tatsächlich der Fall und nach mühsamen Rechnungen, die nicht ausgeführt werden sollen, mit den in (2.2) definierten Größen findet man die folgende Identität (Diese Rechnungen wurden mit den Darstellungen aus L 6.1.f für die Funktionen p_α , q_α , $\alpha = 0(1)q - d_P$ sowie mit den Bezeichnungen aus S 6.3 durchgeführt, da dies einfacher erschien):

$$(*) \quad e_0^T \Psi(\infty) = (e_{q-d_P}^T H_P e_0)^{-1} \left(e_{q-d_P}^T H_P^* - e_{q-d_P}^T H_P \begin{pmatrix} 0 & \dots & 0 \\ H^{-1}(h^*, h^*) \end{pmatrix} \right)$$

1) Der Repräsentant $U_P^{[0]} \in V^>$ unterscheidet sich von $U^{[0]}$ um die um 1 größere Vielfachheit in $t_0 = 0$ (Vgl. (1.5)).

2) Siehe (2.3).

Im zweiten Teil soll eingesehen werden, daß diese Zeile von $\Psi(\infty)$ zusammen mit den anderen, in $\Psi(z)$ z -unabhängigen, gerade die Matrix aus (4.1) ergibt. Wie im Beweis zu S 6.5.a kurz angedeutet wird, kann man

$$(**) \quad H_P^{-1} H_P^* = \begin{pmatrix} e_0^T H_P^{-1} H_P^* \\ H^{-1}(h^*, h^*) \end{pmatrix}$$

einsehen, womit die Zeilen Nummer 1 bis $q-d_P$ der Identität (4.1) verifiziert sind. Mit (*) und (**) folgt für Zeile Nummer 0 :

$$\begin{aligned} e_0^T \Psi(\infty) &= (e_{q-d_P}^T H_P e_0)^{-1} e_{q-d_P}^T \left(H_P^* - H_P (H_P^{-1} H_P^* - \begin{pmatrix} e_0^T H_P^{-1} H_P^* \\ 0 \end{pmatrix}) \right) \\ &= (e_{q-d_P}^T H_P e_0)^{-1} e_{q-d_P}^T H_P \begin{pmatrix} e_0^T H_P^{-1} H_P^* \\ 0 \end{pmatrix} = e_0^T H_P^{-1} H_P^* \end{aligned}$$

Der Beweis zeigt, daß die aufgrund der unterschiedlichen Grade von $q|_0$ einerseits und den Komponenten von q_T^T und p^T andererseits nahe-
liegende Vermutung, Ψ habe zumindest komponentenweise Pole in $z = \infty$, falsch ist.

Wegen der Holomorphie der Matrix Ψ in $z = \infty$ ist bereits klar, daß $Q(\Psi(\infty)) < 1$ A_∞ -Stabilität impliziert. Vor einer genaueren Charakterisierung eine wichtige Bemerkung zur unterschiedlichen Bedeutung dieser Eigenschaft in den Verfahrensklassen $V^{\leq}$ und $V^>$.

4.3 Bemerkung

Die Forderung nach A_∞ -Stabilität kommt aus den Anwendungen der Verfahren auf sogenannte steife Gleichungen. In der stationären Phase der Integration, in der die "steifen Komponenten" der Lösung (eines Differentialgleichungssystems) weitgehend abgeklungen sind, sollen bei einem -an den noch "sichtbaren" Komponenten bemessenen- Vergrößern der Schrittweite die Abgeklungenen nicht wieder (betragsmäßig) ansteigen, sondern weiter fallen. Damit hat man für steife Komponenten in der stationären Phase nurmehr absolute Genauigkeit, aber das genügt. Je kleiner der Wert einer -existierenden- Wachstumsfunktion R in $z = \infty$ betragsmäßig ist, desto stärker werden steife Komponenten auch dann (weiter) gedämpft, wenn verhältnismäßig zu diesen viel zu große Schrittweiten verwendet werden.

Bezogen auf lineare diagonalisierbare Differentialgleichungssysteme

bedeutet A_∞ -Stabilität in der "Situation" (1.6), die hier genau für $V^\leq$ vorliegt, für die bzgl. h hinreichend steifen Komponenten Dämpfung um einen konstanten Faktor in jedem Schritt. Für die Klasse $V^>$, in der die Situation mit (1.11) eine andere ist, gilt zwar bei einem A_∞ -stabilen Repräsentanten (und Anwendung auf (T_μ) , $z := h\mu$) ebenfalls

$$(*) \quad \exists X := X(\infty) \subset \mathbb{K} : \lim_{k \rightarrow \infty} s_k(z) = 0 \quad \forall z \in X - \{\infty\},$$

dennoch liegen andere Verhältnisse als für $V^\leq$ vor. Mittels T 1.2.b und a sowie Formel (2.8) sieht man sofort:

$$(4.2) \quad \begin{cases} R_0(z) = \frac{d_p!}{\beta q!} (-z)^{q-d_p} + O(z^{q-d_p-1}), \quad z \rightarrow \infty, \\ R_k(z) = \frac{e_{q-d_p}^T \Psi^{(\infty)k+1} e_0}{e_{q-d_p}^T \Psi^{(\infty)k} e_0} + O_k(z^{-1}), \quad z \rightarrow \infty, k \in \mathbb{N}. \end{cases}$$

Aus der Darstellung (2.1) folgt für jedes hinreichend kleine $\varepsilon > 0$

$$(**) \quad s_k(z) = \prod_{k \rightarrow \infty} \underbrace{((g(\Psi^{(\infty)})) + \varepsilon)^k}_{< 1} \cdot \prod_{z \rightarrow \infty} (z^{q-d_p}).$$

Formel (4.2), $s_0(z) = 1$, $s_1(z) = R_0(z)$, $z \in \mathbb{C} - N_{q_0}$, und (**) besagen, daß man erst einige Schritte ausführen muß, ehe man in den Bereich der Dämpfung steifer Komponenten gelangt. Der jetzt erklärte Unterschied in der Bedeutung der A_∞ -Stabilität gilt entsprechend auch bzgl. A-Stabilität (im Bereich $|z| \gg 1$) und hat mit einiger Wahrscheinlichkeit kritische Auswirkungen bei der Anwendung der Verfahren aus $V^>$ auf steife Gleichungen. Dieser Unterschied wurde hier aufgrund des Zugangs über die Zwischengrößen R_k (D 2.3.a) gefunden und von anderen Autoren, die sich "nur" mit den Eigenwerten der die Stabilität charakterisierenden Matrix -hier Ψ - beschäftigen, nicht erwähnt (Vgl. Wanner [44], Butcher [3], Burrage/Chipman [2]).

4.4 Satz

Sei $U^{[0]} \in V^>$ mit $q(\Psi(\infty)) < 1$ oder

$$q(\Psi(\infty)) > 1 \text{ und } \infty \in \overline{\mathbb{C}-B} \text{ oder}$$

$$q(\Psi(\infty)) = 1 > |\lambda_1(\infty)| \text{ und } \infty \notin \overline{B}.$$

Dann gilt die folgende Äquivalenz:

$$U^{[0]} \text{ ist } A_\infty\text{-stabil} \iff q(\Psi(\infty)) < 1$$

Bemerkungen

- a) Für den Spezialfall $q=1$, $d_p=0$, $d_1=1$, $i=1(1)p-1$, ist dieses Ergebnis bereits bei Wanner [44], Thm.(11), zu finden. Dort ist $q(\Psi(\infty)) \leq 1$ in Form von 3 Ungleichungen angegeben. Im Fall $q=d_p+1$ ist die obige Aussage (ohne Beweis) bereits in dem Tagungsbericht Fuchs [8], Satz 2, veröffentlicht. Letzteres gilt auch für das anschließende Korollar 4.5.
- b) Sämtliche Voraussetzungen sind nur für " $\implies$ " notwendig, " $\impliedby$ " gilt für alle $U^{[0]} \in V^>$. Im Fall $q=d_p+1$ ist stets $\infty \in \overline{\mathbb{C}-B}$ und $\infty \notin \overline{B}$ (Vgl. S 6.4, #B < ∞).

Beweis

" $\impliedby$ ": Folgt sofort aus S 2.10.c, Bemerkung d zu T 2.2 und Bemerkung a zu D 4.1.

" $\implies$ ": Sei $U^{[0]} \in V^>$ A_∞ -stabil. Annahme: $q(\Psi(\infty)) > 1$.

Aus S 2.10.c folgt die Existenz einer Umgebung $X:=X(\infty) \subset \mathbb{K}$ mit $q(\Psi(z)) > 1$, $z \in X$. Wegen $\infty \in \overline{\mathbb{C}-B}$ gibt es eine Folge $(z_k)_{k \in \mathbb{N}} \subset \mathbb{C}-N_{q|_0} - B$ mit $\lim_{k \rightarrow \infty} z_k = \infty$. Weiter existiert $k_0 \in \mathbb{N}$ mit $z_k \in X, k \geq k_0$.

$$(2.20) \implies |R(z_k)| = |\lambda_0(z_k)| = q(\Psi(z_k)) > 1, k \geq k_0.$$

$$(2.16) \implies (z_k)_{k \geq k_0} \subset \mathbb{C}-R_A \text{ und } \lim_{k \rightarrow \infty} z_k = \infty.$$

Dies steht im widerspruch zur A_∞ -Stabilität. Folglich ist $q(\Psi(\infty)) \leq 1$.

Annahme: $q(\Psi(\infty)) = 1$.

Wegen $|\lambda_0(\infty)| > |\lambda_1(\infty)|$ ergibt sich mit S 2.10.d die Aussage:

Es existiert $X:=X(\infty) \subset \mathbb{K}$ mit X offen, $X \cap N_{q|_0} = \emptyset$ und $|\lambda_0(z)| > |\lambda_1(z)|, z \in X$, und es gilt $\lambda_0 \in \mathcal{X}(X)$. Die "gqT-Voraussetzung" aus S 2.10.d ist hier unnötig -bzw. eingeschränkt auf X automatisch erfüllt-, da sie dazu diene, zu beweisen, daß $\mathbb{K}-N_{q|_0} - A'$ offen ist.

Die Annahme der Existenz einer Umgebung $Y:=Y(\infty) \subset \mathbb{K}$ mit $q(\Psi(z)) < 1$,

$z \in Y$, liefert einen Widerspruch zum Maximumsprinzip, das heißt:

$$(*) \quad \forall Y := Y(\infty) \subset X \quad \exists z \in Y - \{\infty\} : g(Y(z)) \geq 1$$

Wegen $\infty \notin \bar{B}$ existiert $Z := Z(\infty) \subset X$ mit $Z \cap B = \emptyset$; mit (*) folgt: Für alle Umgebungen $Y := Y(\infty) \subset Z \cap X = Z$ existiert ein $z \in Y - \{\infty\}$ derart, daß gemäß (2.20) $|R(z)| = |\lambda_0(z)| = g(Y(z)) \geq 1$ gilt. Darstellung (2.16) von R_A liefert erneut einen Widerspruch zur A_∞ -Stabilität.

4.5 Korollar

- a) Zu beliebigem q, d_p existieren A_∞ -stabile Verfahren.
 b) Sämtliche $U^{[0]} \in V$ sind nicht L-stabil.
 ($U^{[0]} \in V$ heiße L-stabil: $\Leftrightarrow \lim_{z \rightarrow \infty} |R(z)|$ existiert und ist gleich 0.)

Beweis

- a) Für den Fall $q \leq d_p$ siehe (1.9), zu $q > d_p$: Suffizienz von $g(Y(\infty)) < 1$, (4.1), S 6.5.c und Korollar b zu S 6.5.
 b) Bemerkung b zu D 2.6 besagt: $|R(z)| = |\lambda_L(z)|$, $z \in \mathbb{C} - N_{q_0}$.
 L-Stabilität implizierte also $\lambda_{q-d_p}(\infty) = 0$, was der Regularität von $V(\infty)$ widerspräche (Vgl. Korollar zu S 4.2).

4.6 Bemerkung

Da mit (4.1) die explizite Gestalt von $V(\infty)$ gegeben ist, lassen sich etwa bei Variation einer Kollokationsstelle t_j die zugeordneten Parameterintervalle bestimmen, für die A_∞ -Stabilität vorliegt. Dabei geht man wie folgt vor: Die charakteristische Gleichung für die Eigenwerte von $V(\infty)$, nämlich (2.13), $z := \infty$, wird mittels der schon 1956 von Dahlquist [4] im Bereich der Stabilitätstheorie erfolgreich verwendeten Möbiustransformation $\mu \mapsto \lambda(\mu) := (\mu+1)/(\mu-1)$, $\mu \in \mathbb{K}$, in die Form

$$(4.3) \quad g_\infty(\mu) = \sum_{i=0}^{q-d_p+1} b_i(\infty) \mu^i = 0, \quad \mu \neq 1,$$

überführt. Bekanntlich kann man $g(V(\infty)) < 1$ dann mit dem Routh-Kriterium [36] oder Hurwitz-Kriterium [14] (Siehe z.B. Obreschkoff [34]), angewandt auf (4.3), überprüfen, wobei man rationale Ungleichungen in obigem t_j erhält. Sinnvollerweise löst man (numerisch oder analytisch) die entsprechenden Gleichungen und untersucht, in welchen der sich ergebenden t_j -Intervalle $|\lambda_0(\infty)| > 1$

bzw. $|\lambda_0(\infty)| < 1$ vorliegt. Interessant sind natürlich nur diejenigen "Teile" der t_j -Intervalle, in denen man zusätzlich Konvergenz hat. Mit β aus (2.11) sowie

$$(4.4) \quad \Gamma := 2d_p + \sum_{i=1}^{p-1} \frac{d_i}{t_i(1-t_i)}, \quad K := \sum_{i=1}^{p-1} d_i \frac{\left(\frac{1}{2} - t_i\right)}{t_i^2(1-t_i)^2}$$

lautet (2.13), $z := \infty$, (beispielsweise) für $q=d_p+1$ bzw. $q=d_p+2$:

$$(4.5) \quad \lambda(\infty)^2 + \frac{2+\Gamma}{\beta} \lambda(\infty) + \frac{1}{\beta^2} = 0 \quad \text{bzw.}$$

$$(4.6) \quad \lambda(\infty)^3 - \frac{1}{\beta}(3(1+\Gamma) + \frac{1}{2}\Gamma^2 + K)\lambda(\infty)^2 + \frac{1}{\beta^2}(3(1+\Gamma) + \frac{1}{2}\Gamma^2 - K)\lambda(\infty) - \frac{1}{\beta^3} = 0.$$

Die Bestimmung von Parameterintervallen, für die A_∞ -Stabilität vorliegt, sowie eine wichtige Eigenschaft der Größe $|R(\infty)|$ sollen für den Spezialfall $q=d_p+1$ als Vorbereitung für Beispiele noch etwas genauer ausgeführt werden:

4.7 Satz

Sei $U^{[0]} \in V^>$ mit $q = d_p+1$.

a) Sei $j \in \{1, \dots, p-1\}$ fest. Der Deutlichkeit halber werde nachstehend die Abhängigkeit der bisherigen und neu zu definierenden Größen von t_j explizit aufgeschrieben; man setze

$$(4.7) \quad \begin{cases} \varphi_\sigma(t_j) := (1-t_j)^{2d_j} + \beta'^2 t_j^{2d_j} + \sigma \beta'(2+\Gamma') t_j^{d_j} (1-t_j)^{d_j} \\ \quad + \sigma \beta' d_j t_j^{d_j-1} (1-t_j)^{d_j-1}, \\ \bar{\varphi}_\sigma(t_j) := \sigma \beta' t_j^{d_j} - (1-t_j)^{d_j}, \quad \sigma = -1, 1, \text{ mit} \\ \beta' := \prod_{\substack{i=1 \\ i \neq j}}^{p-1} \left(\frac{t_i}{1-t_i} \right)^{d_i}, \quad \Gamma' := 2d_p + \sum_{\substack{i=1 \\ i \neq j}}^{p-1} \frac{d_i}{t_i(1-t_i)}. \end{cases}$$

Die Grenzen zur A_∞ -Stabilität¹⁾ (bzgl. Variation des Parameters t_j in $\mathbb{R}-\{0\}$ ²⁾) sind charakterisiert durch die Vereini-

1) Genauer: Diejenigen Werte t_j^* mit $g(\psi(\infty, t_j^*)) = 1$ (vgl. 4.4).

2) Im Fall $t_j = 1$ werde unter dem Verfahren der (konfluente) Repräsentant $U \in V^<$ mit $U := \lim_{t_j \rightarrow 1} U^{[0]}(t_j)$ verstanden.

gung der reellen, nichtverschwindenden Nullstellen t_j^* von φ_1 , φ_{-1} , für die $|\beta(t_j^*)| \geq 1$ ist, mit denjenigen von $\bar{\varphi}_1$, $\bar{\varphi}_{-1}$, für $\Gamma(t_j^*) \in (-4, 0)$ gilt.

b) Seien $j \in \{1, \dots, p-1\}$ und, in Ergänzung der generellen Voraussetzung, $d_j = 1$ sowie $t_i \in (0, 1)$, $i \in \{1, \dots, p-1\} - \{j\}$. Dann gilt:

$$(4.8) \quad \lim_{t_j \rightarrow 1} \frac{\partial |R(\infty, t_j)|}{\partial t_j} < 0$$

Bemerkung

Gemäß (3.15) und dem nachstehenden Beweis sind die (reellen) Nullstellen t_j^* von $\bar{\varphi}_1$, $\bar{\varphi}_{-1}$ gleichzeitig die Grenzen zur Konvergenz (Genauer: Diejenigen Werte mit $\mathcal{G}(A_2(t_j^*)) = 1$).

Formel (4.8) besagt, daß man die bzgl. A_∞ -Stabilität geeigneteren Verfahren erhält, wenn man genau einen Zwischenknoten t_j aus einer Umgebung rechts von 1 wählt. Besonders interessant ist der Fall $\lim_{t_j \rightarrow 1} |R(\infty, t_j)| = 1$ (Siehe Beispiel 4.8.a). Für die praktische Durchführung ist im Fall $d_p \geq 1$ zu beachten, daß obiges t_j nicht zu nahe bei 1 liegt.

Beweis

a) Man überlegt sich leicht mittels (4.5), daß

$$(*) \quad \lambda_0(\infty, t_j) = \frac{1}{\beta(t_j)} \frac{1 \pm \sqrt{1 + 4/\Gamma(t_j)}}{1 \mp \sqrt{1 + 4/\Gamma(t_j)}}, \quad \Gamma(t_j) \in [-4, 0], \text{ und}$$

$$|\lambda_0(\infty, t_j)| = |\beta(t_j)|^{-1}, \quad \Gamma(t_j) \in [-4, 0],$$

gelten. Die Eigenwerte von $\Psi(\infty, t_j)$ sind entweder beide reell oder beide komplex und genau für $\Gamma(t_j) \in (-4, 0)$ komplex. Die Polynome φ_σ erhält man durch Multiplikation der Gleichung (4.5) mit $\beta^2 t_j^{2d_j} (\neq 0)$ und $\lambda(\infty) := (-1)^{m-d_p} \sigma$ (reeller Fall), wobei $|\lambda_1(\infty, t_j)| \leq 1$ nach Vieta $|\beta(t_j)| \geq 1$ erzwingt.

Im komplexen Fall führt die Umrechnung von $1/\beta(t_j) = (-1)^{m-d_p} \sigma$ auf $\bar{\varphi}_\sigma(t_j) = 0$. Die Äquivalenz (Charakterisierung) ergibt sich aus einfachen Überlegungen. Man beachte, daß (auch) im Grenzfall $\lim_{t_j \rightarrow 1} |\lambda_0(\infty, t_j)| = 1$ $t_j = 1$ Nullstelle von φ_1 oder φ_{-1} ist.

b) Durch Taylorentwicklung von $R(\infty, t_j) = \lambda_0(\infty, t_j)$ um $t_j = 1$ (Vgl. (*)) erhält man die Grenzwerte

$$\lim_{t_j \rightarrow 1} R(\infty, t_j) = (-1)^{m-q} (\beta')^{-1},$$

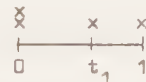
$$\lim_{t_j \rightarrow 1} \frac{\partial}{\partial t_j} R(\infty, t_j) = (-1)^{m-d_p} (\beta')^{-1} (4 + \Gamma'), \text{ also}$$

$$\lim_{t_j \rightarrow 1} \frac{\partial}{\partial t_j} |R(\infty, t_j)| = -|\beta'|^{-1} (4 + \Gamma').$$

Mit der Voraussetzung folgt $\Gamma' > 0 > -4$.

4.8 Beispiele

a) $p=q=2$, $d_2=d_1=1$, $t_1 > 0$



Konvergenz liegt genau für $t_1 \geq 1/2$ vor. S 4.4, (4.7) zeigen:

- Das Verfahren ist genau für $t_1 \in (1, (1+\sqrt{5})/2)$ A_∞ -stabil.

Die Skizze für die beiden Funktionen $t_1 \mapsto |R(\infty, t_1)|$ und $t_1 \mapsto g(A_2(t_1))$ ist die gleiche wie in Beispiel 3.5.b, nur daß dort $1/|\beta|$ durch $g(A_2(t_1))$ und $g(A_2)$ durch $|R(\infty, t_1)|$ zu ersetzen sind. Dieser Zusammenhang ist keineswegs zufällig und wird im folgenden Abschnitt 4-2 noch allgemein formuliert und bewiesen.

Hier tritt der in der Bemerkung zu S 4.7 angesprochene Fall ein, wo der rechte Randpunkt 1 des "aktuellen Integrationsintervalls" für t_1 Grenze zur A_∞ -Stabilität ist. Die weiter unten angesprochenen Computerzeichnungen ("Plots") zeigen, daß in $R_A(t_1)$ qualitativ ein wesentlicher Unterschied zwischen $t_1 \geq 1$ und $t_1 \leq 1$ besteht, was den hiesigen Fall anbetrifft.

Diese Zeichnungen wurden mit einem Programm erstellt, daß in wesentlichen Teilen von Kühle [22] stammt. Im Programm werden die Höhenlinien $c=1$ des Betrages einer Funktion $R: \mathbb{C} \rightarrow \mathbb{K}$ ausgehend von $z=0$ mit "einem" Newtonverfahren "verfolgt". Da man im allgemeinen nur einen Zweig der Höhenlinien von $|R(\cdot)|$ erhält, sind die Zeichnungen nur als Hilfsmittel anzusehen. In das Programm von Kühle wurden als Funktionen für $q \leq d_p$ R aus (1.7), im Fall $q=d_p+1$ -statt R aus D 2.6- $Q(\Psi(\cdot))$ implementiert. Laut (2.16), D 2.7 und S 6.4 gehen dabei die endlich vielen Punkte

$$R_A = \{z \in \mathbb{C} \mid Q(\Psi(z)) < 1\}$$

verloren, die jedoch praktisch unbrauchbar sind.

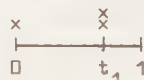
In den Plots sind die Ränder ∂R_A der Stabilitätsregionen für verschiedene Parameter t_1 zu sehen, wobei die Region R_A selbst diejenige der beiden Zusammenhangskomponenten ¹⁾ von $K - \partial R_A$ ist, die "in $z = 0$ links liegt". Die Plots finden sich im Abbildungsteil. Nun zurück zum Beispiel:

Plot 1,2 : Grenze zur A_∞ -Stabilität in $t_1^\infty := 1$, qualitativer Unterschied zwischen $t_1 < 1$ und $t_1 > 1$ (Im Fall $t_1 = 1$ ist $\partial R_A = i\mathbb{R}$).

Plot 3 : Grenze zur A_∞ -Stabilität in $t_2^\infty := (1 + \sqrt{5})/2 \doteq 1.618$.

Die Zeichnungen legen die Vermutung nahe, daß eine Grenze zur A_∞ -Stabilität auch Grenze zur A-Stabilität sein kann (Hier $t_1^\infty = 1$; $t_2^\infty = (1 + \sqrt{5})/2$ jedoch nicht).

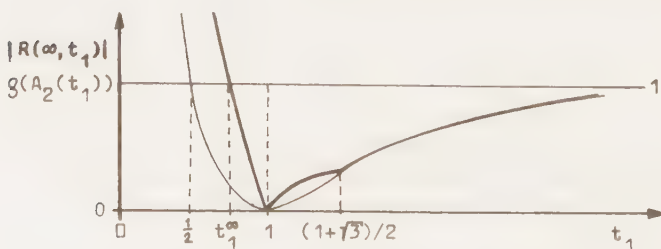
b) $p=2, q=1, d_2=0; d_1=2, t_1 > 0$



Konvergenz liegt genau für $t_1 \geq 1/2$ vor. Aus S 4.4, (4.7) folgt:

- Das Verfahren ist genau für $t_1 \in (t_1^\infty, \infty)$ mit $t_1^\infty = \frac{1}{2}(1 + \sqrt{1/3}) \doteq 0.7887$ A_∞ -stabil.

Zu den Verhältnissen bei diesem Beispiel die folgende Skizze:



Plot 4,5,6 : Numerische Bestimmung von t_1^∞ mittels Intervallschachtelung mit dem Ergebnis: $0.788 < t_1^\infty < 0.789$.

■ 4-1

¹⁾ Im allgemeinen ist es keineswegs selbstverständlich, daß $K - \partial R_A$ aus nur zwei Zusammenhangskomponenten besteht (Vgl. z.B. Grigorieff [11], S.220, Fuchs [7], S.61, Jeltsch/Nevalinna [21], Figur 1).

4-2 Zusammenhang zwischen A_∞ -Stabilität und Konvergenz

Die folgende Aussage beinhaltet einen unerwarteten Zusammenhang zwischen A_∞ -Stabilität und Konvergenz:

4.9 Theorem

Sei $U^{[0]} \in V$ mit $q < d_p$ oder

$$q \geq d_p^{(1)} \text{ und } g(\Psi(\infty)) < 1 \text{ oder}$$

$$q \geq d_p^{(1)} \text{ und } g(\Psi(\infty)) > 1 \text{ und } \omega \in \overline{\mathbb{C}-B} \text{ und } t_i \in (0,1),$$

$i=1(1)p-1$. Ferner sei wie in S 4.2 $U_p^{[0]} := (0, U^{[0]})$. Dann folgt :

$$U^{[0]} \text{ ist } A_\infty\text{-stabil} \iff U_p^{[0]} \text{ ist konvergent}$$

Bemerkung

Die Aussage gilt insbesondere an der "Nahtstelle", wo $U^{[0]} \in V^\leq$, $U_p^{[0]} \in V^>$ ist.

Beweis

(i) $q < d_p$: Es gilt $U_p^{[0]} \in V^\leq$, und sämtliche Verfahren aus $V^\leq$ sind konvergent. $U^{[0]}$ ist gemäß (1.9) A_∞ -stabil.

(ii) $q = d_p$: Aus Formel (2.3) in Mülthei [29] bzw. aus (1.9) folgt $|R(\infty)| = 1/|\beta|$. Wegen $R \in \mathcal{R}(K-P_R)$, des Maximumsprinzips und D 4.1 folgt die Äquivalenz: $U^{[0]}$ ist genau dann A_∞ -stabil, wenn $1/|\beta| < 1$ gilt; $q_p = d_p + 1$, Beispiel a zu S 6.5 und S 6.5.c zeigen $g(A_{2,p}) = 1/|\beta|$. In allen betrachteten Fällen gilt:

$$(*) \quad U_p^{[0]} \text{ ist konvergent} \iff \underline{g(A_{2,p}) < 1}$$

(iii) $q > d_p$: Die Aussage (*) gilt aufgrund der Voraussetzungen und wegen Satz 5 aus Mülthei [26] auch hier; entsprechend folgt mittels S 4.4 die Äquivalenz:

$$(**) \quad U^{[0]} \text{ ist } A_\infty\text{-stabil} \iff \underline{g(\Psi(\infty)) < 1}$$

Der wesentliche -und für (*) unter $q > d_p$ schon benutzte- Schritt des Beweises, der jetzt -noch einmal- anzuwenden ist, nämlich der Nachweis der Gleichheit von $g(\Psi(\infty))$ und $g(A_{2,p})$ im allgemeinen Fall $q > d_p$ folgt unmittelbar aus (4.1) und S 6.5.c. ●

¹⁾ Im Fall $q = d_p$ ist $\Psi(z) = (R(z))$ mit R aus (1.7) und $B := \emptyset$.

Mit T 4.9 ist der erwähnte Zusammenhang zwischen Beispiel 3.6.b ($U_{(p)}^{[0]} = (0,0,0,t_1,1)$) und Beispiel 4.8.a ($U^{[0]} = (0,0,t_1,1)$) klar.

Weiter kann man Parameterintervalle, für die Konvergenz vorliegt, wie folgt bestimmen:

- a) Der Fall $q=d_p+1$ ist mit (3.15) klar.
- b) Im Fall $q=d_p+2$ gilt S 4.7.a entsprechend.
- c) Für $q > d_p+2$ kann man wie in B 4.6 beschrieben vorgehen.

5 A-Stabilität und Stabilitätsregionen

5.1 Definition

Sei $U^{[0]} \in V$. $U^{[0]}$ heiße A-stabil : $\Leftrightarrow R_A \supset \mathbb{C}^- := \{z \in \mathbb{C} \mid \operatorname{Re}(z) < 0\}$ ●

Die Vermutung, es könne wegen des Pols der Funktion R_0 aus (1.11) in $z = \infty$ (Vgl. T 1.2.b) für $q > d_p$ keine A-stabilen Verfahren geben, ist nach den Ergebnissen aus Kapitel 4, insbesondere K 4.5.a, nicht aufrecht zu erhalten. Die Existenz A_∞ -stabiler Verfahren $U^{[0]} \in V^>$ (Siehe auch Plot 2, 4 und 6) legt die Vermutung der Existenz A-stabiler Verfahren in $V^>$ nahe.

Ziel des folgenden und abschließenden Kapitels soll es sein, das Vorhandensein A-stabiler Verfahren in $V^>$ "sehr allgemein" zu beweisen und mit Beispielen zu belegen, wie man konstruktiv Parameterkonfigurationen bestimmen kann, für die die gewünschte Eigenschaft vorliegt. Auf eine nähere Ausführung der Konstruktionsmethode, die aus dem Existenzbeweis direkt hervorgeht und wie dieser auf Abschätzungen beruht, wird verzichtet, um den Umfang der vorliegenden Arbeit nicht zu groß werden zu lassen.

Beim Beweis der Existenz A-stabiler Verfahren und bei der Konstruktion gehen die beiden Eigenschaften (2.20) und (2.21) von R wesentlich ein.

Es sei (noch einmal) an der Bedeutung der A-Stabilität bei der praktischen Anwendung der Verfahren, vor allem auf steife Differentialgleichungssysteme¹⁾, einerseits, andererseits aber auch an die in § 4.3 angesprochene Problematik erinnert.

1) Eine Zusammenstellung von Literatur zur numerischen Behandlung solcher Systeme findet sich in Fuchs [7], S.11 .

5-1 Existenz A-stabiler Verfahren (Konfluenz)

Die Idee auf dem Weg zum Beweis der Existenz A-stabiler Verfahren in $V^>$ ist, sich bzgl. des Parametersatzes $U^{[0]} \in V^>$ "in der Nähe" eines A-stabilen Verfahrens aus $V^<$ zu bewegen, was gleich genauer expliziert wird. Die Ausführung dieser "Stetigkeitsidee" erweist sich als komplizierter, als man zunächst vermuten könnte.

Andererseits sind die Resultate dann aber besser, als zu erwarten, d.h. man kann sich -bei geeigneter Wahl des Grenzübergangs- mit dem Parametersatz $U^{[0]} \in V^>$ relativ "weit von dem aus $V^<$ entfernen" und trotzdem A-Stabilität "halten" (Siehe Tabelle in 8 5.13).

5.2 Bemerkung (und Definition)

Man betrachte eine von einem reellen Parameter δ abhängige Schar von Verfahren $(U^{[0]}(\delta))_{\delta \in \mathbb{R} - \{0\} \subset V^>}$, die die Eigenschaft besitzen, daß dasjenige Verfahren, welches dem vermöge

$$\tilde{U} := \tilde{U}^{[0]} := \lim_{\delta \rightarrow 0} U^{[0]}(\delta)$$

definierten Parametersatz zugeordnet ist, in der Klasse $V^<$ liegt. Man betrachte hierzu (1.5). Diese Eigenschaft beim Grenzübergang $\delta \rightarrow 0$ soll im folgenden Konfluenz (in $V^<$) genannt werden. Konfluenz ist demnach der bzgl. des Parametersatzes stetige Übergang einer Verfahrensschar aus $V^>$ in ein Verfahren aus $V^<$ (Diese Idee findet sich auch in Korollar b zu S 6.5 und wurde bereits in K 4.5.a verwendet).

Zur Kennzeichnung werden im folgenden sämtliche $U^{[0]}(\delta)$ zugeordneten Größen, die von δ abhängen, mit der (ggf. zusätzlichen) Variablen δ sowie sämtliche $\tilde{U}$ zugeordneten Größen mit $\sim$ versehen. Beispiele: $\Psi(\cdot, \delta)$ bezeichne die $U^{[0]}(\delta)$ gemäß T 2.2 zugeordnete Matrix, $\tilde{R} = \tilde{P}/\tilde{Q}$ ($\tilde{P}(0) := \tilde{Q}(0) := 1$) sei die $\tilde{U}$ gemäß (1.7) bzw. T 1.2.b zugeordnete Wachstumsfunktion. Der Begriff der Konfluenz muß für die theoretische Behandlung formal strikter gefaßt und soll der Einfachheit halber noch etwas eingeschränkt werden:

$$(5.1) \quad \left\{ \begin{array}{l} \text{Sei } \text{ggT}(\tilde{P}, \tilde{Q}) = 1. \\ \text{Es existiere } K \neq \emptyset \text{ mit } K \subset \{1, \dots, p-1\}, \text{ so daß} \\ \text{gilt: Für } j \in K \text{ ist } t_j \text{ stetig von } \delta \text{ abhängig,} \\ \text{für } j \in \{1, \dots, p-1\} - K \text{ von } \delta \text{ unabhängig. Die übrigen} \\ \text{Parameter seien für hinreichend kleines } |\delta| \neq 0 \\ \text{von } \delta \text{ unabhängig. Es gelte } 1 - t_j(\delta) = O(\delta), \delta \rightarrow 0, j \in K. \\ \text{Sei } d^{(K)} := \sum_{j \in K} d_j \gg q - d_p. \end{array} \right.$$

Desweiteren werde die folgende, von (1.5) etwas abweichende Schreibweise für die Parametersätze verwendet:

$$(5.2) \quad \left\{ \begin{array}{l} U^{[0]}(\delta) = (\underbrace{0, \dots, 0}_q, \underbrace{(t_1, \dots, t_i)}_{d_1} \underset{i \notin K}, \underbrace{(t_j(\delta), \dots, t_j(\delta))}_{d_j} \underset{j \in K}, \underbrace{1, \dots, 1}_{d_p}), \\ \tilde{U} = (\underbrace{0, \dots, 0}_q, \underbrace{(t_1, \dots, t_i)}_{d_1} \underset{i \notin K}, \underbrace{1, \dots, 1}_{\tilde{d}_p = d_p + d(K)}) \end{array} \right.$$

Es ist zu vermerken, daß "ein großer Teil" der Verfahren aus $V^>$ "in der Nähe" von $V^{\leq}$ liegt, genauer lassen sich sämtliche Verfahren $U^{[0]} \in V^>$, die der schwachen Nebenbedingung

$$(5.3) \quad \sum_{i=1}^{p-1} d_i \geq q - d_p$$

genügen, in der oben beschriebenen Weise konfluent in ein Verfahren aus $V^{\leq}$ überführen.

Das vorgehen zum beweis der Existenz A -stabiler verfahren in $V^>$ ist nun das folgende: Vorgegeben wird ein A -stabiles Verfahren $\tilde{U} \in V^{\leq}$ mit den zusätzlichen Eigenschaften $\partial \tilde{R}_A \subset \mathbb{C}^+ \cup \{0\}$ und $|\tilde{R}(\infty)| < 1$. Hierzu wird eine Schar von verfahren $(U^{[0]}(\delta))_{\delta \neq 0} \subset V^>$ gemäß (5.1), (5.2) konstruiert.

Die eben erwähnten Eigenschaften besagen, daß der Rand $\partial \tilde{R}_A$ der Stabilitätsregion bis auf den Nullpunkt ganz in $\mathbb{C}^+$ verläuft und sich nicht asymptotisch der imaginären Achse nähert. In $V^{\leq}$ existieren Verfahren, die diesen Bedingungen genügen. Für spätere Beispiele sind diejenigen $\tilde{U} \in V^{\leq}$, denen als Wachstumsfunktionen $\tilde{R}$ subdiagonale Padéapproximationen zugeordnet sind, besonders einfach, da man insbesondere nach den Untersuchungen von Wanner/Hairer/Nørsett [45] die Funktionen $\tilde{R}$ auf $\mathbb{R}$ genau kennt, was später (Vgl. L 5.4) wichtig ist.

Es besteht die Hoffnung, daß die "Nähe" der Parametersätze (5.2) für hinreichend kleines $|\delta|$ eine "Ähnlichkeit" der Stabilitätsregionen $\tilde{R}_A$ und $R_A(\delta)$ bewirkt, was keineswegs trivial ist, sich aber bestätigen wird. Plot 7 (siehe Abbildungsteil) verdeutlicht diese Idee. Bzgl. einer einelementigen Menge K in (5.1) siehe auch zweite Skizze in B 1.4.a.

Der wesentliche Gedanke für den Beweis selbst ist sehr einfach zu formulieren, jedoch technisch nur mit Mühe durchführbar:

Der betragsgrößte Eigenwert $\lambda_0(.,\delta)$ der Matrix $\Psi(.,\delta)$ geht für $\delta \rightarrow 0$ in die $\tilde{U}$ zugeordnete Wachstumsfunktion $\tilde{R}$ über, während die anderen Eigenwerte $\lambda_i(.,\delta)$, $i=1(1)q-d_0$, bei diesem Grenzübergang gegen 0 konvergieren. Letzteres gilt -in einer Umgebung von $z=0$ - auch für $x_r(z,\delta)$. Hierauf kommt man bei genauerem Betrachten von (2.4) bzw. (6.2) und (2.18) unter Berücksichtigung von S 6.5.c und Korollar b zu S 6.5.

Wegen (2.16), S 2.10.a und D 5.1 ist für die angestrebte A-Stabilität in $V^>$ die Existenz eines $\delta \neq 0$ mit $|\lambda_0(z,\delta)| < 1$, $z \in \mathbb{C}^-$ zu zeigen. Naheliegenderweise wird bei der hiesigen Vorgehensweise ein $\bar{\delta} > 0$ mit

$$|\lambda_0(z,\delta)| < 1, \quad z \in \mathbb{C}^-, \quad 0 < |\delta| \leq \bar{\delta},$$

konstruiert. Natürlich benutzt man zum einen das Maximumsprinzip, so daß im wesentlichen noch die Existenz eines $\bar{\delta} > 0$ mit

$$|\lambda_0(z,\delta)| \leq 1, \quad z \in \overline{\mathbb{R}}, \quad 0 < |\delta| \leq \bar{\delta},$$

zu beweisen bleibt, und zum anderen -wegen der obigen Konvergenzaussagen und der Kenntnis der Funktion $\tilde{R}$ - die Dreiecksungleichung

$$|\lambda_0(z,\delta)| \leq |\lambda_0(z,\delta) - \tilde{R}(z)| + |\tilde{R}(z)|.$$

Man könnte meinen, daß nun wegen der Konvergenz von $\lambda_0(.,\delta)$ gegen $\tilde{R}$ alles klar ist, es gibt jedoch noch zwei wesentliche Probleme: Zum einen gilt $\tilde{R}(0) = 1$, so daß man sich die Differenz $\lambda_0(z,\delta) - \tilde{R}(z)$ sowie $\tilde{R}(z)$ in einer Umgebung von $z=0$ genauer ansehen muß. Dieses Problem wird mittels (2.21) bzw. (3.12) überwunden, was die Untersuchung der Vereinigung $F(\delta)$ der Ausnahmемengen erfordert.

Zum anderen benötigt man für den Existenzbeweis eine gleichmäßige Abschätzung von $\lambda_0(.,\delta) - \tilde{R}$ auf $\overline{\mathbb{R}}$, für die darüber hinausgehende Konstruktion A-stabiler Verfahrensscharen in $V^>$ kommt es natürlich auch darauf an, mit welcher Ordnung in δ die obige Differenz gegen Null strebt. Da die dazu notwendigen Hilfsmittel L 5.3, L 5.6.a(ii), c und K 5.7.b,c kompliziert erscheinen mögen, sei vermerkt, daß eine gewisse Vereinfachung dieser möglich wäre, wenn man nur auf die gleichmäßige Abschätzbarkeit der vorkommenden Größen Wert legte, die genaueren Aussagen der Lemmata jedoch einen wichtigen Hinweis für die Konstruktion liefern (Gemeint ist die geeignete Anordnung der Kollokationsstellen t_1 "um die 1", vgl. K 5.9, Beispiele 5.10 und 5.12).

Das folgende Lemma wird im Beweis zu L 5.6 verwendet:

5.3 Lemma

Gegeben seien $R \in \mathbb{C} \setminus \{0\}$, $k, l, n \in \mathbb{N}$, $\delta_0 \in \mathbb{R}_+$ sowie Funktionen μ_i , $i=0(1)n$, und c_i , $i=0(1)n-1$, : $[-\delta_0, \delta_0] \rightarrow \mathbb{C}$ mit

$$\begin{aligned}\mu_0(\delta) &= R + O(\delta^l), \mu_i(\delta) = O(\delta^k), \delta \rightarrow 0, |\delta| \leq \delta_0, i=1(1)n, \\ c_i(\delta) &= O(\delta^{(n-i)k}), \delta \rightarrow 0, |\delta| \leq \delta_0, i=0(1)n-1.\end{aligned}$$

Für die Nullstellen $\lambda_i(\delta)$, $i=0(1)n$, des Polynoms $p(., \lambda) \in \Pi_{n+1}$, definiert vermöge

$$(5.4) \quad p(\lambda, \delta) := \prod_{i=0}^n (\lambda - \mu_i(\delta)) + \sum_{i=0}^{n-1} c_i(\delta) \lambda^i, \lambda \in \mathbb{C},$$

gilt für ein geeignetes δ_1 mit $0 < \delta_1 \leq \delta_0$:

$$\lambda_0(\delta) = \mu_0(\delta) + O(\delta^k), \lambda_i(\delta) = O(\delta^k), \delta \rightarrow 0, |\delta| \leq \delta_1, i=1(1)n.$$

Hierbei seien die Nullstellen von $p(., \delta)$ -entsprechend dem in der Einleitung zu Abschnitt 2-2 Gesagten- dem Betrage nach fallend angeordnet.

Beweis

Wegen $R \neq 0$ ist die übliche Konditionsanalyse für einfache Nullstellen bzgl. $\mu_0(\delta)$ möglich und ergibt die Aussage über die Größe $\lambda_0(\delta)$. Dividiert man den Linearfaktor $\lambda - \lambda_0(\delta)$ aus (5.4) ab, so erhält man

$$\sigma_1(\lambda_1(\delta), \dots, \lambda_n(\delta)) = O(\delta^{1k}), \delta \rightarrow 0, i=n(-1)1.$$

Der Rest der Behauptung folgt nun mit Ostrowski [35], S.276.

5.4 Lemma

Sei $U \in V^k$. U sei A -stabil mit $\partial R_A \subset \mathbb{C}^+ U \setminus \{0\}$ und $|R(\infty)| < 1$. Gemäß (1.5) sei m die Anzahl der Parameter von U . Dann gilt:

- a) $\exists \varepsilon, K_0 > 0$: $|R(iy)| \leq 1 - K_0 |y|^{m+1}$, $y \in \mathbb{R}$, $|y| \leq \varepsilon$.
 b) $\forall \varepsilon' > 0 \exists K_1 < 1$: $|R(iy)| \leq K_1$, $y \in \mathbb{R}$, $|y| \geq \varepsilon'$.

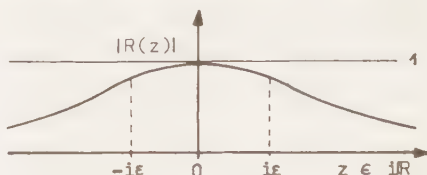
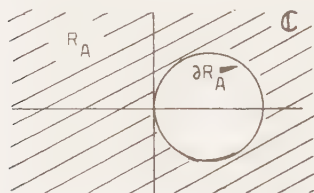
Bemerkung

Zu jedem Paar $(j, k) \in \mathbb{N}^2$ mit $j \geq k$ gibt es mindestens ein $U =: U^{(j, k)} \in V^k$, dessen Wachstumsfunktion $R =: R^{(j, k)}$ die Padéapproximation vom

Typ (j,k) ($R^{(j,k)} = P_k/Q_j$, $\text{grad } P_k = k$, $\text{grad } Q_j = j$) ist (Mülthei [24], Satz 6 und [29], Satz 5). Bekanntlich besitzen die $U^{(j,k)}$ mit $j=k+1$ oder $j=k+2$ die in L 5.4 geforderten Eigenschaften.

Beweis

Hier sei nur die Idee angedeutet: Verfahren, die den Voraussetzungen genügen, besitzen eine Stabilitätsregion R_A bzw. eine Wachstumsfunktion R mit den wie folgt skizzierten Eigenschaften:



a) Mit T 1.2.b,a folgt die Entwicklung

$$e^z R(z) = cz^{\mu+1} + c'z^{\mu+2} + O(z^{\mu+3}), z \rightarrow 0, c \neq 0, \mu > m.$$

Man berechne $|R(iy)|^2$ und unterscheide die Fälle μ gerade bzw. μ ungerade.

b) Trivial wegen der Stetigkeit von R auf $i\mathbb{R}$ und $|R(iy)| < 1, \epsilon' \leq |y| \leq \infty$.

5.5 Bemerkung

Sei $U^{[0]} \in V^>$ mit $\text{ggT}(q_0, q_0 a_{q-d_p}, \dots, q_0 a_0) = 1$ (Vgl. (2.14)). Die durch die algebraische Gleichung (2.13) definierte algebraische Funktion λ ist auf der zugeordneten Riemannschen Fläche $\mathcal{R}_\lambda$ holomorph bis auf ihre Pole N_{q_0} (Vgl. Beweis zu S 2.10.d). Das Maximumsprinzip für Riemannsche Flächen besagt:

$$(5.5) \quad \left\{ \begin{array}{l} 1. \quad N_{q_0} \cap \overline{\mathbb{C}^-} = \emptyset, \\ 2. \quad |\lambda_0(iy)| \leq 1, y \in \mathbb{R}, \end{array} \right.$$

ist äquivalent mit der Aussage

$$|\lambda_0(z)| \leq 1, z \in \overline{\mathbb{C}^-}, |\lambda_0(z)| < 1, z \in \mathbb{C}^-.$$

Bemerkung d zu T 2.2 bzw. (2.16) und S 2.10.a besagen hiermit:

(5.5) impliziert A-Stabilität von $U^{[0]}$. Zur Idee, das Maximumsprinzip bei Verfahren "mit (3.4)" anzuwenden, vgl. z.B. Genin [10], S.387.

Das Verhalten der Ausnahmemengen aus D 2.7 für $\delta \rightarrow 0$ ist wichtig für den konkreten Nachweis von (5.5) (Vgl. auch Bemerkung c zu D 2.7). Daher:

5.6 Lemma

Gegeben seien eine Schar von Verfahren $(U^{(0)}(\delta))_{\delta \neq 0} \subset V^>$ sowie $\tilde{U}$ gemäß (5.1), (5.2).

a) Für die vom Konfluenzparameter δ abhängigen Ausnahmemengen, die $U^{(0)}(\delta)$ zugeordnet sind, gelten die folgenden Aussagen:

$$(i) \quad \lim_{\delta \rightarrow 0} N_{q_0}(\cdot, \delta) = N_{\tilde{q}}.$$

(ii) Zu jeder Umgebung $V := \bigcup_{z \in N_{\tilde{q}} \cup N_{\tilde{q}}} V(z)$ existiert ein $\delta_A > 0$ mit

$$A'(\delta) \cap (K-V) = \emptyset, \quad 0 < |\delta| \leq \delta_A.$$

(iii) Es existieren $\varepsilon, \delta_F > 0$ mit $F(\delta) \cap B_{\varepsilon}(0)^1 = \emptyset, \quad 0 < |\delta| \leq \delta_F.$

b) Die Bedingung (2.14) -aufgeschrieben mit der δ -Abhängigkeit der Größen $q_0, a_i, i=0(1)q-d_p$, -ist für hinreichend kleines $|\delta| (\neq 0)$ erfüllt (Vgl. B 5.5).

c) Es existieren $\varepsilon_x, \delta_x > 0$, so daß die Abschätzung

$$x_r(z, \delta) = O(\delta^{d^{(k)} - (q-d_p-1)}), \quad \delta \rightarrow 0, \quad 0 < |\delta| \leq \delta_x,$$

gleichmäßig bzgl. $z \in B_{\varepsilon_x}(0)$ gilt.

Beweis

a) (i) : Trivial.

(ii) : Für das Folgende sei $k := d^{(k)} - (q-d_p-1).$

Aus dem Beweis von Korollar b zu S 6.5 folgt mit S 6.5.c

$$(*) \quad H(\delta)^{-1}(H^*(\delta), H^*(\delta)) = O(\delta^k), \quad \delta \rightarrow 0.$$

Es sei gemäß (6.2)

$$\mu_0(z, \delta) := \nu_{00}(z, \delta) = (p_0(z, \delta) - q_1(z, \delta) H(\delta)^{-1} H^*(\delta)) / q_0(z, \delta),$$

$\mu_i(\delta), i=1(1)q-d_p$, bezeichne die Eigenwerte von $H^{-1}(\delta) H^*(\delta).$

¹⁾ $B_{\varepsilon}(z_0) := \{z \in \mathbb{C} \mid |z - z_0| < \varepsilon\}.$

Mittels der charakteristischen Gleichung (2.4), Anwendung der Cramerschen Regel bzgl. $\det(\lambda I_{\mathbb{R}} - H(\delta))^{-1} H^*(\delta) (\lambda I_{\mathbb{R}} - H(\delta))^{-1} H^*(\delta))^{-1}$, (*) sowie der Taylorentwicklung von $\mu_0(\cdot, \delta)$ um $\delta = 0$ lassen sich die folgenden Aussagen beweisen:

$$(**) \left\{ \begin{array}{l} \text{Für jede Umgebung } W := \bigcup_{z \in N_{\tilde{R}} \cap N_{\tilde{Q}}} Y(z) \text{ und für jede Umgebung} \\ Y := \bigcup_{z \in N_{\tilde{R}} \cup N_{\tilde{Q}}} Y(z) \text{ mit } Y \supset W \text{ existiert ein} \\ \delta_W > 0, \quad \text{so daß mit } n := q - d_p \text{ gilt:} \\ \\ f_{\psi(z, \delta)}(\lambda) = (\lambda - \mu_0(z, \delta)) \prod_{i=1}^n (\lambda - \mu_i(\delta)) + \sum_{i=0}^{n-1} c_i(z, \delta) \lambda^i, \text{ mit} \\ \mu_0(z, \delta) = \tilde{R}(z) + O(\delta), \mu_i(\delta) = O(\delta^k), \delta \rightarrow 0, |\delta| \leq \delta_W^{-1}, i=1(1)n, \\ c_i(z, \delta) = O(\delta^{(n-i)k}), \delta \rightarrow 0, |\delta| \leq \delta_W^{-1}, i=0(1)n-1, \\ \\ \text{gleichmäßig bzgl. } z \in K-W. \text{ Weiter existieren } \underline{x}, \bar{x} \\ \text{mit } 0 < \underline{x} \leq |\tilde{R}(z)| \leq \bar{x} < \infty, z \in K-Y. \end{array} \right.$$

Unter Berücksichtigung der Gleichmäßigkeit der obigen Fehlerausagen liegen in (**) die Voraussetzungen von L 5.3 mit $l:=1, \delta_0:=\delta_W$ und $p(\cdot, \delta) := f_{\psi(z, \delta)}(\cdot)$ vor. Anwendung dieses Hilfssatzes ergibt mit $K-Y \subset K-W$ die Existenz eines $\delta_1 > 0$, so daß gilt:

$$(5.6) \quad \left\{ \begin{array}{l} \lambda_0(z, \delta) = \gamma_{00}(z, \delta) + O(\delta^k) = \tilde{R}(z) + O(\delta^1), \\ \lambda_i(z, \delta) = O(\delta^k), \delta \rightarrow 0, |\delta| \leq \delta_1^{-1}, i=1(1)q-d_p, \\ \\ \text{gleichmäßig bzgl. } z \in K-Y. \end{array} \right.$$

Aus (5.1) folgt $k \geq 1$. Die Behauptung (ii) ergibt sich aus der Definition von $A^1(\delta)$ sowie der gleichmäßigen Beschränktheit "von Null weg" der Funktion $|\tilde{R}(\cdot)|$ auf $K-Y$.

(iii): Folgt mittels (*) bzw. Korollar b zu S 6.5 analog wie im Beweis zu L 3.3.b. Dabei ist zu überlegen, daß der Durchschnitt über alle sich (dort) ergebenden Mengen $\tilde{W}_\delta$ bzw. W_δ die Null noch als inneren Punkt enthält, was daran liegt, daß $g(A_2(\delta))$ für hinreichend kleines $|\delta|$ gleichmäßig kleiner als 1 ist.

1) In $\delta = 0$ sind die δ -abhängigen Größen stetig ergänzt.

b)

$$\begin{aligned} a_{q-d_p}(\cdot, \delta) q_0(\cdot, \delta) &= -q_0(\cdot, \delta) \operatorname{sp} \Psi(\cdot, \delta) = \\ &= -p_0(\cdot, \delta) + q_1(\cdot, \delta)^T H(\delta)^{-1} H^*(\delta) - q_0(\cdot, \delta) \operatorname{sp}(H(\delta)^{-1} H^*(\delta)). \end{aligned}$$

Für hinreichend kleines $|\delta|$ gilt $m - \tilde{d}_p \leq \operatorname{grad}(a_{q-d_p}(\cdot, \delta) q_0(\cdot, \delta)) \leq m - q$. Mittels (*) folgt aus der obigen Formel, daß von den Nullstellen von $a_{q-d_p}(\cdot, \delta) q_0(\cdot, \delta)$ für $\delta \rightarrow 0$ $m - \tilde{d}_p$ stetig in diejenigen von $\tilde{P}$ übergehen, die ggf. restlichen streben nach ∞ . Hier sei noch einmal an $\tilde{d}_p = d_p + d^{(k)} \gg q$ erinnert. Die Behauptung folgt nun mit a(i) und der Voraussetzung $\operatorname{ggT}(\tilde{P}, \tilde{Q}) = 1$, die in (5.1) enthalten ist.

c) Folgt aus a(iii), (2.18), $R(0, \delta) = 1$, S 2.10.a, c mittels (*).

5.7 Korollar

Es gelten die Voraussetzungen von L 5.6 und es sei $k := \underline{d^{(k)}} - (q - d_p - 1)$. Dann folgt:

a) Es existieren $\varepsilon, \delta_0 > 0$ mit

$$\lambda_0(z, \delta) = R(z, \delta) = \tilde{g}(z, U(z, \delta), m), \quad 0 < |\delta| \leq \delta_0, \quad z \in B_\varepsilon(0).$$

b) Zu jeder Umgebung $Y := \bigcup_{z \in N_{\tilde{R}} \cup N_{\tilde{Q}}} Y(z)$ existiert ein $\delta_1 > 0$ mit

$$\lambda_0(z, \delta) = \tilde{R}(z) + O(\delta^1), \quad \lambda_i(z, \delta) = O(\delta^k), \quad i = 1(1)q - d_p, \delta \rightarrow 0,$$

$0 < |\delta| \leq \delta_1$, gleichmäßig bzgl. $z \in \mathbb{K} - Y$.

c) Im Fall $k > 1$ (d.h. $\tilde{R}(\infty) = 0$) gilt: Zu jedem $K_\infty > 0$ existieren eine Umgebung $X(\infty)$ und ein $\delta_\infty > 0$ mit

$$|\lambda_0(z, \delta) - \tilde{R}(z)| \leq K_\infty^{-1}, \quad 0 < |\delta| \leq \delta_\infty, \quad z \in X(\infty).$$

1) Eine b entsprechende Abschätzung bzgl. $X(\infty)$, in der K_∞ durch $K_\infty |\delta|^1$ zu ersetzen wäre, läßt sich noch für gewisse Beispiele mit $q = d_p + 1$ zeigen. In allgemeineren Fällen sind solche Abschätzungen wegen der möglichen Verzweigung in zwei Variablen sicher nur schwer zu erhalten, andererseits höchstens von theoretischem Interesse, da für das "Ziel (5.5)2." eine Umgebung $X(\infty)$ für $k > 1$ unproblematisch ist, wie eben die Aussage in c zeigt.

Beweis

- a) L 5.6.a(iii), S 2.10.a, (2.21) .
 b) Formel (5.6) bzw. L 5.6.a(ii) .
 c) Für $k > 1$ gilt laut (5.2) $\tilde{q} < \tilde{d}_p$, folglich wegen (1.9) $\infty \in N_{\tilde{R}}$.
 Man überlegt sich mit (5.1), (6.2) und (4.1), daß $\mathcal{W}(\cdot, \cdot)$ und somit auch $\lambda_0(\cdot, \cdot)$ in einer Umgebung von $(\infty, 0)$ stetig von (z, δ) abhängt. Aus $\lambda_0(\infty, 0) = 0$ folgt die Behauptung. ●

Mit L 5.4, B 5.5, L 5.6 und K 5.7 sind nun die Hilfsmittel bereitgestellt, um den angestrebten Hauptsatz problemlos beweisen zu können. Er lautet:

5.8 Theorem

Gegeben sei $\tilde{U} \in V^*$ mit $\tilde{q}, \tilde{d}_p \in \mathbb{N}$, $\tilde{P}, \tilde{Q}$ seien teilerfremd und $\tilde{P}$ besitze keine imaginären Nullstellen; $\tilde{U}$ sei A-stabil mit $\partial \tilde{R}_A \subset \mathbb{C}^+ \cup \{0\}$ und $|\tilde{R}(\infty)| < 1$. Zu beliebigem $d^{(K)} \in \{\tilde{d}_p - \tilde{q} + 1, \dots, \tilde{d}_p\}$ konstruiere man eine Schar von Verfahren $(U^{[0]}(\delta))_{\delta \neq 0} \subset V$ gemäß (5.1), (5.2). Dann existiert ein $\bar{\delta} > 0$, so daß die folgende Aussage gilt:

$$U^{[0]}(\delta) \text{ ist A-stabil für } 0 < |\delta| \leq \bar{\delta} .$$

Bemerkung

Wie allgemein ist die Aussage von T 5.8 ? Die "Art der Konfluenz" ist nicht wesentlich eingeschränkt; mit $\tilde{q} = 0$ oder $\tilde{d}_p = 0$ "läßt sich keine Konfluenz machen". Die Voraussetzungen bzgl. $\tilde{P}$ und $\tilde{Q}$ entstehen aus formaler Bequemlichkeit.

Bedeutend als Restriktion sind lediglich $\partial \tilde{R}_A \subset \mathbb{C}^+ \cup \{0\}$, $|\tilde{R}(\infty)| < 1$. Bei Konfluenz in $\tilde{U} \in V^*$ (beispielsweise) mit $\partial \tilde{R}_A = \mathbb{IR}$ versagt die vorgestellte Abschätzungsmethode. Hier zeigen Zeichnungen der zugeordneten Stabilitätsregionen, daß der folgende naheliegende Fall auftreten kann: Es existiert ein $\bar{\delta} > 0$, so daß $U^{[0]}(\delta)$ A-stabil für $0 < \delta \leq \bar{\delta}$, nicht A-stabil für $0 < -\delta \leq \bar{\delta}$ ist (Siehe Plot 1,2, $\delta := t_1 - 1$).

Beweis

B 5.5 und L 5.6.b besagen, daß für die $U^{[0]}(\delta)$ zugeordneten Größen noch das Folgende zu zeigen ist: Es existiert ein $\bar{\delta} > 0$ mit

$$1. N_{q_0}(\cdot, \delta) \cap \overline{\mathbb{C}} = \emptyset, \quad 2. |\lambda_0(iy, \delta)| \leq 1, y \in \overline{\mathbb{R}}, 0 < |\delta| \leq \delta.$$

Aus der A-Stabilität von $\tilde{U}$ und der Teilerfremdheit von $\tilde{P}$, $\tilde{Q}$ folgt $N_{\tilde{Q}} \cap \overline{\mathbb{C}} = \emptyset$, was mit L 5.6.a(i) die Existenz eines $\delta_q > 0$ mit

$$N_{q_0}(\cdot, \delta) \cap \overline{\mathbb{C}} = \emptyset, \quad 0 < |\delta| \leq \delta_q,$$

impliziert. Die Abschätzung von $|\lambda_0(\cdot, \delta)|$ über $\overline{\mathbb{R}}$ wird in drei Teilen vorgenommen:

- (i) In einer Umgebung $\overline{B_\varepsilon(0)} \cap \overline{\mathbb{R}} = [-i\varepsilon, i\varepsilon]$,
- (ii) In einer Umgebung $\overline{X_\eta(\infty)} \cap \overline{\mathbb{R}} = \{z \in \overline{\mathbb{R}} \mid |z| \geq \eta\}$,
- (iii) Im restlichen Bereich $[-i\eta, -i\varepsilon] \cup [i\varepsilon, i\eta]$.

Zu (i): (Hier geht die Darstellung (2.21) über (3.12) wesentlich ein. Dies gilt insbesondere auch bzgl. konkreter Beispiele, wofür die unter (*) folgenden Abschätzungen explizit durchzuführen sind.) Wegen K 5.7.a, L 5.6.c, (2.2) und (2.19) existieren Größen ε' , δ'_0 , ε_q , $\tilde{\varepsilon}_q$, $\tilde{\gamma}_q$, $\gamma_x > 0$ derart, daß die folgenden Aussagen gelten:

$$(*) \quad \begin{cases} \lambda_0(z, \delta) = R(z, \delta) = \tilde{G}(z, U(z, \delta), m), \quad \|x_T(z, \delta)\| \leq \gamma_x |\delta|^k, \\ \varepsilon_q \leq |\tilde{Q}(z)| \leq \tilde{\varepsilon}_q, \quad \varepsilon_q \leq |Q(z, \delta)|, \\ |q_0(z, \delta) - \tilde{q}(z)| \leq \gamma_0 |\delta|^1, \quad 0 < |\delta| \leq \delta'_0, \quad z \in B_{\varepsilon'}(0) \end{cases}$$

Trivialerweise gilt (im gleichen Bereich bzgl. (z, δ)) die Ungleichung

$$(5.7) \quad |\lambda_0(z, \delta)| \leq |(e^z - \tilde{R}(z)) - (e^z - R(z, \delta))| + |\tilde{R}(z)|.$$

Die beiden Identitäten in T 1.2.a und L 3.3.a -(3.12)- liefern nach einfachen Umformungen in (5.7) mit (*) die Existenz eines $K_0 > 0$ mit

$$|\lambda_0(iy, \delta)| \leq K_0 |\delta| |y|^{m+1} + |\tilde{R}(iy)|, \quad 0 < |\delta| \leq \delta'_0, \quad |y| \leq \varepsilon'.$$

Aus L 5.4.a folgt schließlich: $\exists \varepsilon, \delta_0 > 0 : |\lambda_0(iy, \delta)| \leq 1$, $0 < |\delta| \leq \delta_0$, $|y| \leq \varepsilon$.

Zu (ii): Wegen $|\tilde{R}(\infty)| < 1$ und K 5.7.b bzw. $\tilde{R}(\infty) = 0$ und K 5.7.c gilt die gewünschte Abschätzung in einer hinreichend "kleinen" Umgebung von ∞ , d.h. insbesondere: $\exists \eta > \varepsilon, \delta_\infty > 0 : |\lambda_0(iy, \delta)| \leq 1$, $0 < |\delta| \leq \delta_\infty$, $|y| \geq \eta$.

Zu (iii): Aufgrund der Voraussetzungen bzgl. $\tilde{P}$ und $\tilde{Q}$ sowie der A-Stabilität von $\tilde{U}$ besitzen weder $\tilde{R}$ noch $\tilde{Q}$ imaginäre Nullstellen.

Folglich ist in K 5.7.b eine Wahl von γ mit $(K-\gamma) \cap \mathbb{R} \supset [-i\gamma, i\gamma]$ möglich. K 5.7.b und L 5.4.b zeigen: $\exists \delta_1 > 0 : |\lambda_0(i\gamma, \delta)| \leq 1$, $0 < |\delta| \leq \delta_1$, $\varepsilon \leq |\gamma| \leq \eta$.

Mit $\bar{\delta} := \min(\delta_0, \delta_1, \delta_\infty)$ folgt die Behauptung.

Mit dem Beweis zu T 5.8 läßt sich das Folgende einsehen: Es existieren $\varepsilon, \delta_0, \delta_1', K_0, \tilde{K}_0, K_1 > 0$, $\tilde{K}_1 < 1$ mit

$$(5.8) \quad \begin{cases} |\lambda_0(i\gamma, \delta)| \leq K_0 |\delta| \cdot |\gamma|^{m+1} + 1 - \tilde{K}_0 |\gamma|^{m+1}, & |\gamma| \leq \varepsilon, 0 < |\delta| \leq \delta_0' \\ |\lambda_0(i\gamma, \delta)| \leq K_1 |\delta| + \tilde{K}_1^{-1}, & |\gamma| \geq \varepsilon, 0 < |\delta| \leq \delta_1' \end{cases}$$

Das folgende Korollar beinhaltet einen wichtigen Gedanken für die Verbesserung von Parameterbereichen mit A-Stabilität bei konkreten Beispielen.

5.9 Korollar

Es gelte für ein $\sigma \in \mathbb{N}$ mit $2 \leq \sigma \leq d^{(K)}$:

$$(5.9) \quad \sigma_i \left(\underbrace{(1-t_j(\delta), \dots, 1-t_j(\delta))}_{d_j} \right)_{j \in K} = O(\delta^\sigma), \delta \rightarrow 0, i=1(1)d^{(K)}$$

Dann läßt sich in (5.8) $|\delta|$ durch $|\delta|^\sigma$ ersetzen.

Bemerkung

Wegen der zu erfüllenden Forderung (5.5)2. ist es plausibel, daß die höheren Potenzen in (5.8) einen größeren δ -Bereich "bringen", für den A-Stabilität vorliegt, auch wenn man hier nichts über die Abhängigkeit der Konstanten K_0, K_1 von σ aussagen kann. Daß diese Idee zu einer wesentlichen Verbesserung führt, werden die Beispiele

1) Der Term $K_1 |\delta|$ rührt hierbei im Grunde von der Abschätzung von $|\lambda_0(i\gamma, \delta) - \tilde{R}(i\gamma)|$ her. Die angegebene Ungleichung ist gemäß K 5.7.c auch im Fall $\tilde{R}(\infty) = 0$ in einer Umgebung von ∞ ($|\gamma| \geq \eta$) korrekt, wenn auch künstlich um den δ -abhängigen Summanden erweitert. Man beachte hierzu das in Fußnote 1) zu K 5.7 Gesagte.

im folgenden Abschnitt 5-2 demonstrieren.

Beweis

Idee: Die Bedingungen in (5.9) bewirken, daß in $\psi(z, \delta)$ δ nur in Form δ^σ oder $\delta^{\sigma'}$, $\sigma' \geq \sigma$, auftritt (Vgl. (6.2), (6.3), (2.2), L 6.1.f). Daß sich diese Tatsache bis in die Abschätzungen (5.8) "fortsetzt", muß man insbesondere über die zu modifizierenden "Stationen" (***) im Beweis zu L 5.6 und (*) im Beweis zu T 5.8 nachweisen.

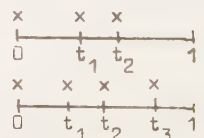
5-2 Beispiele A-stabiler Verfahren

Bevor die im zurückliegenden Abschnitt angegebene Abschätzungsmethode auf Beispiele angewendet wird, einige Bemerkungen zu denjenigen in der Literatur vorhandenen Resultaten über Stabilitätstheorie bzgl. der Testgleichung (T_μ) , die so allgemein sind, daß man sie zumindest auf Teile der Klasse $V^>$ anwenden kann. In diesem Zusammenhang sind zwei wesentliche Richtungen zu unterscheiden:

Zum einen gibt es Arbeiten, die sich mit "linearer Stabilitätstheorie" für solche Verfahren beschäftigen, denen bei Anwendung auf (T_μ) eine charakteristische Relation der Form (3.4) zugeordnet ist (Für die Verfahren aus der Klasse V von Mülthei gilt (3.4) mit $M = m - q$, $K = 1 + \max(0, q - d_p)$). Hier ist in erster Linie Jeltsch zu nennen, der sich etwa in [15-19] und [6] (mit Friedli) sowie in [20-21] (jeweils mit Nevanlinna) mit zahlreichen Aspekten der linearen Stabilitätstheorie eingehend beschäftigt (Vgl. Einleitung). A-Stabilität selbst spielt in diesen Arbeiten eine untergeordnete Rolle, dagegen finden sich in [6], [16], [18-19] Kriterien für A_0 -Stabilität und "steife Stabilität" (Definitionen siehe [16], S.761), wobei für die spezielle Klasse von "Brown's MSMD-methods" praktisch leicht überprüfbare Kriterien erzielt werden. Vorher hat schon Genin in [10], S.387, das Maximumsprinzip (5.5) für Verfahren "mit" (3.4) aufgeschrieben. Letztere Arbeit beschäftigt sich weiter mit der Verallgemeinerung der klassischen Frage von Dahlquist 1963, [5], Thm.2.2, nämlich welche maximale Ordnung man zu gegebenen Zahlen M, K in (3.4) unter der Nebenbedingung A-Stabilität erreichen kann. In den oben genannten Arbeiten wird nicht konkret auf den Zusammenhang zu Kollokationsmethoden eingegangen.

In einer zweiten Gruppe von Arbeiten wird dieser Zusammenhang, insbesondere die Angabe von Parameterkonfigurationen, für die A-Stabilität vorliegt, untersucht. Hierzu gibt es jedoch bisher nur vereinzelt Resultate. Die einzigen beiden Fälle aus der Klasse $V^>$, die in den nachfolgend angegebenen Arbeiten theoretisch auf A-Stabilität untersucht wurden, sind:

$$(5.10) \quad \left\{ \begin{array}{l} p=3, q=1, d_3=0; d_1=d_2=1, \\ p=4, q=1, d_4=0; d_1=d_2=d_3=1. \end{array} \right.$$



Die Ergebnisse finden sich in einer Arbeit von Butcher [3] über verallgemeinerte singly-implizite RK-Methoden, welche eine Teilklasse von $V^>$ bilden, deren Charakterisierung in Abschnitt 1-1 angegeben

wurde. Insbesondere sind in (5.10) die Kollokationsstellen in der von "singly-implicitness" erzwungenen Weise gekoppelt. Die Resultate bzgl. A-Stabilität stammen aus einer unveröffentlichten Arbeit von Wanner [44] und basieren auf der Order-Star-Technik von Wanner/Hairer/Nørsett [45]. Diese ist im "algebraischen Fall" ((3.4) mit $K > 1$) allerdings nur äußerst knapp beschrieben, in [44] finden sich, was die Überprüfung der A-Stabilität anbelangt, nurmehr Skizzen. Die erwähnte Technik führt zu Charakterisierungsaussagen, in [44] wird jedoch davon Gebrauch gemacht, daß bei der oben angedeuteten Kopplung der Kollokationsknoten das zugeordnete "Wannerpolynom" q_0 nur eine reelle $(m-1)$ -fache Nullstelle besitzt.

Die in der vorliegenden Arbeit eingeschlagene Vorgehensweise führt mit T 5.8 zwar nur zu Teilmengen von den tatsächlichen Parameterbereichen, für die A-Stabilität vorliegt, ist jedoch elementar -wenn auch mühsam- auf allgemeinere Fälle anwendbar. Anschließend an die Einleitung dieses Abschnittes soll insbesondere auf die Beispiele (5.10), jedoch mit einer anderen Kopplung der Knoten, eingegangen werden. In einer weiteren Arbeit, [2] von Burrage/Chipman, findet sich eine größere Menge von verallgemeinerten singly-impliciten Verfahren auf A-Stabilität untersucht, wobei die Ergebnisse allerdings numerisch durch "Abtasten" von $q(\psi(.))$ über $\mathbb{C}^-$ erzielt werden.

Dabei zeigt sich ein wesentliches Problem bei A-Stabilität in Verbindung mit singly-implicitness: Die Kollokationsstellen liegen zum Teil weit außerhalb des "aktuellen Integrationsintervalls", und das schon im zweiteinfachsten Fall $p=3, q=1$, was von den beiden Autoren nur am Rande erwähnt wird ([2], S.4 oben). Daher ist es hier ein weiteres Ziel, A-Stabilität mit in der Nähe und innerhalb des (aktuellen) Integrationsintervalls liegenden Zwischenknoten zu erreichen. Die Bemerkung zu S 4.7 legt die Vermutung nahe, daß dies erleichtert wird, wenn ein Knoten "rechts" von 1 liegt.

Um für konkrete Beispiele die in T 5.8 mit $\bar{\delta}$ bezeichnete Größe zu berechnen, sind zahlreiche Abschätzungen von rationalen Funktionen in den zwei Variablen ¹⁾ (z, δ) zu erstellen. Auf dem (mühsamen) Weg gelangt man über (5.7) und (5.6) zu Abschätzungen der Form (5.8), aus denen man δ_0 und δ_1 (Vgl. Beweis zu T 5.8) und somit $\bar{\delta}$ elementar berechnet. Um den Umfang der vorliegenden Arbeit nicht zu groß werden zu lassen, soll, wie gesagt, auf die nähere Explizierung dieser Kon-

¹⁾ Die Rationalität dieser Funktionen gilt natürlich nur bei rationaler Abhängigkeit der $t_j(\delta)$, $j \in K$, von δ (Vgl. auch (5.13)).

struktionsmethode verzichtet werden. Die Idee der Abschätzungen beruht auf den Darstellungen (2.20), (2.21) ¹⁾ sowie zwei Ungleichungen, die hier (noch einmal) erwähnt seien:

$$(5.11) \quad |\lambda_0(z, \delta)| = |R(z, \delta)| \leq \underbrace{|(e^z - \tilde{R}(z)) - (e^z - R(z, \delta))|}_{\text{Approximationseigenschaft der Funktionen } \tilde{R}, R(., \delta)} + \underbrace{|\tilde{R}(z)|}_{\text{Vorgabe L5.4, Bsp.: Padé-Approx. an die Exponentialfunktion } \tilde{R}: \text{Nørsett/Wanner, T1.2, [32] Wanner/Haire; } R(., \delta): \text{Fuchs, L3.3, [8] (q=d, 1) Nørsett [45]}}$$

$z \in F(\delta), (2.20)$
 $z \in \overline{\mathbb{R}}, |z| \leq \varepsilon$

$$(5.12) \quad |R(z, \delta)| \leq |\lambda_0(z, \delta)| \leq \underbrace{|\lambda_0(z, \delta) - \psi_{00}(z, \delta)|}_{\text{Konditionsproblem L5.3, vgl. (5.6) (Ausnahme } \tilde{R}(\infty)=0, \text{ Umgebung von } z=\infty)} + \underbrace{|\psi_{00}(z, \delta) - \tilde{R}(z)|}_{\text{Taylorentwicklung Vgl. (**) im Beweis}} + |\tilde{R}(z)|$$

$S \text{ 2.10.a}$
 $z \in \overline{\mathbb{R}}, |z| \geq \varepsilon$

Noch ganz kurz und vereinfacht zur Vorgehensweise bei den folgenden Beispielen:

0. Absichern der technischen Voraussetzungen, nämlich der ggT-Bedingung (L 5.6.b) einerseits, Bestimmung von $\varepsilon, \delta_T > 0$ mit $F(\delta) \cap [-i\varepsilon, i\varepsilon] = \emptyset, A'(\delta) \cap \mathbb{R} = \emptyset, 0 < |\delta| \leq \delta_T$, andererseits.
1. Erster Teil des Maximumsprinzips: Bestimmung eines $\delta_q > 0$ mit $N_{q|_0}(., \delta) \cap \overline{\mathbb{C}} = \emptyset, 0 < |\delta| \leq \delta_q$.
2. Zweiter Teil des Maximumsprinzips: Bestimmung eines δ'_0 bzw. δ'_1 mit $0 < \delta'_0 \leq \min(\delta_T, \delta_q)$ bzw. $0 < \delta'_1 \leq \min(\delta_T, \delta_q)$ derart, daß die erste bzw. zweite Zeile in (5.8) gilt (Man beachte K 5.9).
3. Berechnung von

$$\bar{\delta} := \min \left(\delta'_0, \sqrt[{\sigma}]{K'_0/K_0}, \delta'_1, \sqrt[{\sigma}]{(1-\tilde{R}_1)/K_1} \right).$$

¹⁾ Man beachte, daß bei den Abschätzungen im Bereich $\{z \in \overline{\mathbb{R}}, |z| \leq \varepsilon\}$ (3.12) und damit (2.21) eingeht. Hierbei wird $\|x_r(., \delta)\|$ mittels (2.18) abgeschätzt (Vgl. Beweis zu L 5.6.c).

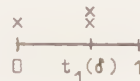
5.10 Beispiele

Im folgenden seien $q=d_p+1$, $k:=d^{(h)}$ (Anzahl der Konfluenzknoten einschließlich Vielfachheiten) und σ die zu gegebener Schar $(U^{[0]}(\delta))_{\delta \neq 0}$ in K 5.9 maximal mögliche Größe gleichen Namens. Einfachste Konfluenzmöglichkeit gemäß T 5.8 ist der Übergang in das Verfahren, dem als Wachstumsfunktion die subdiagonale Padéapproximation vom Typ (2,1) zugeordnet ist. Hierbei gibt es im wesentlichen zwei unterschiedliche Möglichkeiten:

a) $\tilde{U} = (0, 1, 1)$, $U^{[0]}(\delta) := (0, 1-\delta, 1-\delta)$, $-\infty < \delta < 1^1$ bzw.

$p=2$, $q=1$, $d_2=0$; $d_1=2^1$, $t_1(\delta)=1-\delta$

Hier gilt $k=2$, $\sigma=1$.



(i) Konvergenz(ordnung)

Die folgenden Größen werden gemäß (3.15) berechnet. Mit $c_{m+1}(\delta)$ werde der Koeffizient von z^{m+1} in der sich aus (3.12) ergebenden Entwicklung bezeichnet.

$$c_{m+1}(\delta) = \frac{1}{6} \int_0^1 (\tau - (1-\delta))^2 \left(\tau - \frac{\delta^2}{2\delta-1} \right) d\tau, \text{ falls}$$

$$g(A_2(\delta)) = \delta^2 / (1-\delta)^2 < 1.$$

Es folgt:

$$v(\delta) = \begin{cases} -1, & \delta > 1/2, \\ 2, & \delta = 1/2, \\ 3, & \delta < 1/2. \end{cases}$$

Inbesondere existiert kein superkonvergenzparameter d.h. kein δ^S mit $v(\delta^S) = m+1 = 4$ bzw. $c_{m+1}(\delta^S) = 0$, $g(A_2(\delta^S)) \leq 1$, $A_2(\delta^S) \neq (1)$.

(ii) A-Stabilität (numerisch)

Der Bereich der A-Stabilität wurde zunächst numerisch -durch Zeichnung der Ränder der Stabilitätsregionen $R_{\lambda}(\delta)$ - bestimmt, siehe Plot 4.5.6 (Vgl. auch Beispiel 4.8.b). Es ergibt sich A-Stabilität für

$$-\infty < \delta \leq 0.211^{(2)}.$$

¹⁾ Im folgenden sei stets $U^{[0]}(\delta) := \tilde{U}$ (die Angabe der festen Parameter bezieht sich hierbei gemäß (5.1) auf den Fall hinreichend kleines $|\delta| \neq 0$).

²⁾ Untersuchung der Grenzen der Parameterbereiche mit A-Stabilität auf 3 Stellen.

(iii) A-Stabilität (laut I 5.8, (5.11), (5.12))

Die Abschätzungen (5.8) wurden mit $\varepsilon := 2$, $\delta'_0 := \delta'_1 := 0.1$ erstellt¹⁾ und es folgte $\bar{\delta} = \min(0.1, 0.0561, 0.1, 0.0511)$, was beweist:

$U^{[0]}(\delta) = (0, 1 - \delta, 1 - \delta)$ ist A-stabil, falls $|\delta| \leq 0.0511$ gilt.

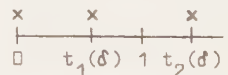
Da bei den Abschätzungen Vorzeichen nicht berücksichtigt werden, wäre als maximaler A-Stabilitätsbereich " $|\delta| \leq 0.211$ " beweisbar. Insofern wurde der "tatsächliche" Bereich der A-Stabilität um einen Faktor ≈ 4 unterschätzt.

Im folgenden Beispiel - zweite Möglichkeit der Konfluenz in $(0, 1, 1)$ - soll die Idee aus K 5.9 angewendet werden, um zu einer Verbesserung zu kommen. Die Knoten sollen so angeordnet werden, daß $\sigma = 2$ bzw. $R(z, \delta) - \tilde{R}(z) = O_z(\delta^2)$ gilt. Es ist zu vermuten, daß bei einer solchen Wahl der Knoten die Region $R_A(\delta)$ in der Nähe von $\delta = 0$ weniger stark von δ abhängt als in a, mit anderen Worten "länger" in der Nähe der Region $\tilde{R}_A$ der $(2, 1)$ -Padéapproximation liegt. Anders formuliert ist es wegen (5.5)2. plausibel, daß $|\delta|^2$ statt $|\delta|$ in (5.8) - bei Einschränkung auf bzgl. $\delta = 0$ symmetrische Parameterintervalle - einen größeren Bereich mit A-Stabilität bringt. Diese naheliegende Vermutung erweist sich als richtig.

b) $\tilde{U} = (0, 1, 1)$, $U^{[0]}(\delta) := (0, 1 - \delta, 1 + \delta)$, $|\delta| < 1$, bzw.

$p=3$, $q=1$, $d_3=0$; $d_1=d_2=1$, $t_1(\delta)=1-\delta$, $t_2(\delta)=1+\delta$

Es ist $k = \sigma = 2$.



Gegenüber a besteht zunächst der praktische Vorteil, daß keine Vielfachheit d_j größer als 1 ist (Vgl. (1.3)).

(i) Konvergenz(ordnung)

$$c_{m+1}(\delta) = \frac{1}{6} \int_0^1 ((\tau-1)^2 - \delta^2)(\tau - \delta^2) d\tau, \quad \text{falls}$$

$$q(\Lambda_\sigma(\delta)) = \delta^2 / |1 - \delta^2| \leq 1.$$

$$\nu(\delta) = \begin{cases} -1 & , \quad |\delta| > 1/\sqrt{2} \quad , \\ 3 & , \quad |\delta| \leq 1/\sqrt{2} \quad , \quad |\delta| \neq \delta^S \quad , \\ 4 & , \quad |\delta| = \delta^S := \sqrt{(5 - \sqrt{13})/12} \approx 0.340887166 \end{cases}$$

¹⁾ In Verfeinerung von (5.8) wurden auch $|\delta|^2$ -Terme berücksichtigt und "für" (5.5)2. quadratische Ungleichungen aufgelöst.

Für $|\delta| > 1/\sqrt{2}$ gilt $g(A_2(\delta)) > 1$. Hierbei ist Divergenz unterstellt oder, anders gesagt, γ vermöge $\gamma := -1$, sofern $g(A_2) > 1$, zum Teil neu definiert, da sich Satz 5 aus Mülthei [26] nicht anwenden läßt. Im Gegensatz zu Beispiel a existiert hier nach den obigen Angaben als weiterer Vorteil ein Superkonvergenzparameter.

(ii) A-Stabilität (numerisch)

Die Zeichnungen Plot 8,9,10 bestätigen die oben geäußerte Vermutung. A-Stabilität liegt (offensichtlich) für

$$|\delta| \leq 0.330$$

vor. Der Superkonvergenzparameter liegt zwar nicht mehr in diesem Bereich, jedoch noch im δ -Intervall der A_∞ -Stabilität (Näheres dazu im abschließenden Abschnitt 5-3).

(iii) A-Stabilität (theoretisch)

Mit $\varepsilon := 2$, $\delta_0' := 0.2$, $\delta_1' := 0.15$ folgt $\bar{\delta} = \min(0.2, 0.1528, 0.15, 0.1307)$, was besagt:

$U^{(0)}(\delta) = (0, 1 - \delta, 1 + \delta)$ ist A-stabil, sofern $|\delta| \leq 0.1307$ gilt.

Bei Einschränkung auf bzgl. $\delta = 0$ symmetrische Parameterintervalle ist gegenüber a also nicht nur der tatsächliche A-Stabilitätsbereich größer, sondern auch der Unterschätzfaktor durch die Abschätzungsmethode, der hier ≈ 2.5 beträgt, kleiner, was jeweils auf $\sigma = 2$ an Stelle von $\sigma = 1$ zurückzuführen ist.

Für das noch folgende Beispiel soll der A-Stabilitätsbereich nurmehr numerisch bestimmt werden. Nach dem Fortschritt im Beispiel 5.10.b gegenüber a ist es konsequent, die Möglichkeit $\sigma \geq 3$ in K 5.9 anzuviesieren. Der Einfachheit halber sei für das noch folgende

$$(5.13) \quad 1 - t_j(\delta) = c_j \delta \quad \text{mit} \quad c_j \in \mathbb{R} - \{0\}, \quad j \in K.$$

Das Resultat der näheren Untersuchung obiger Möglichkeit ist jedoch (zunächst) negativ:

5.11 Lemma

Seien $q=d_p+1$, $k=d^{(k)} \geq 2$, $\sigma \in \mathbb{N}$, $2 \leq \sigma \leq k$, $C := ((c_j, \dots, c_j)_{j \in K}) \in \mathbb{R}^k$.

a) Unter der Einschränkung (5.13) gilt (5.9)

"mit σ " genau dann, wenn das System nichtlinearer (Un)gleichungen

$$(5.14) \quad \sigma_i(C)=0, i=1(1)\sigma-1, \quad \sigma_k(C) \neq 0$$

mindestens eine Lösung $C \in \mathbb{R}^k$ besitzt.

b) In (5.14) ist für alle $k \geq 2$ bei Beschränkung auf reelle Lösungen maximal $\sigma = 2$ möglich.

c)

$$\frac{1}{k} \sum_{j \in K} d_j t_j(\delta) = 1, \delta \in \mathbb{R} - \{0\} \iff \sigma \geq 2$$

Bemerkungen

- a) Beispiel 5.10.b belegt, daß sich eine geeignete Anordnung der Konfluenzkollokationsknoten um die 1 über den Parameter σ günstig für den δ -Bereich der A-Stabilität von $U^{[0,1]}(\delta)$ auswirkt. Wären komplexe Kollokationsstellen zulässig, so wäre die Anordnung "Einheitswurzeln" auf dem Kreis vom Radius δ um 1 optimal und bewirkt $\sigma = k$.
- b) Aussage c besagt, daß die Differenz $R(z, \delta) - \tilde{R}(z)$ genau dann von der Fehlerordnung $O_2(\delta^2)$ ist, wenn die Konfluenzkollokationsstellen den Schwerpunkt 1 besitzen.

Beweis

a), c) Trivial.

b) Man beweist durch Induktion die folgende Aussage:

Für alle $k \in \mathbb{N}$, $k \geq 3$, $b_{k-3}, \dots, b_0 \in \mathbb{R}$ mit $b_j \neq 0$ für ein $j \in \{0, \dots, k-3\}$ besitzt das Polynom $p_k \in \Pi_k$, definiert durch

$$p_k(c) := c^k + b_{k-3}c^{k-3} + \dots + b_0, \quad c \in \mathbb{C},$$

weniger als k reelle Nullstellen.

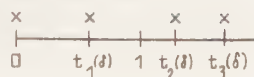
Im Fall $\sigma \geq 2$ ist $R(z, \delta) - \tilde{R}(z) = O_2(\delta^2)$. Wäre $\sigma_1(C) = \sigma_2(C) = 0$ mit $C \in \mathbb{R}^k$ möglich, so wäre diese Differenz von der Fehlerordnung $O_2(\delta^3)$. Insofern ist es plausibel, die Größe $\sigma_2(C)$, die für $\sigma=2$ $k-2$ freie Parameter enthält, in geeigneter Weise zu minimieren. Dies geschieht im folgenden Beispiel mit $k=3$. Dabei wurden die c_j , $j \in K$, so gewählt, daß zwar $|\sigma_2(C)|$ "nicht ganz minimal", dafür aber $d_j=1$, $j \in K$, ist.

5.12 Beispiel

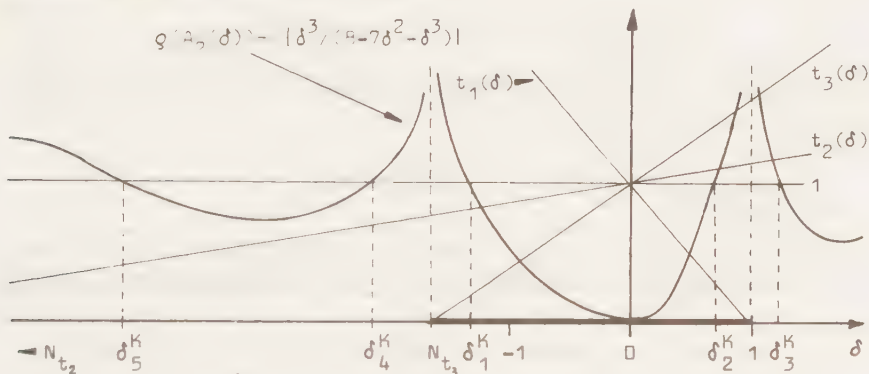
$\tilde{U} = (0, 1, 1, 1)$, $U^{(0)}(\delta) := (0, 1-\delta, 1+\frac{1}{2}(1-\frac{1}{\sqrt{2}})\delta, 1+\frac{1}{2}(1+\frac{1}{\sqrt{2}})\delta)$, d.h.

$p=4$, $q=1$, $d_4=0$; $d_1=d_2=d_3=1$.

Es ist $k=3$, $\sigma=2$.



(i) Konvergenzbereich und -ordnung



$$\delta_1^K = -\sqrt{8/7} \approx -1.069044967, \quad \delta_2^K \approx 0.9482747641,$$

$$\delta_3^K = +\sqrt{8/7}, \quad \delta_4^K \approx -1.370556459, \quad \delta_5^K \approx -3.077718305.$$

Die Größen δ_1^K , $i=2,4,5$ wurden mit einem Programm zur simultanen Berechnung von Polynomnullstellen von Werner [46] bestimmt. Für das Folgende sei $I := [\delta_5^K, \delta_4^K] \cup (\delta_1^K, \delta_2^K) \cup (\delta_3^K, \infty)$. Für $q(A_2(\delta)) < 1$ gilt

$$c_{m+1}(\delta) = \frac{1}{10(41)^2(8-7\delta^2)} (-96+364\delta^2+100\delta^3-245\delta^4-210\delta^5-30\delta^6), \Rightarrow$$

$$v(\delta) = \begin{cases} -1, & \delta \in \mathbb{R} \setminus I \\ 3, & |\delta| = \sqrt{8/7} \\ 4, & \delta \in I \setminus \{\delta_1^S, \delta_2^S\}, \quad \text{mit } \delta_1^S \approx -0.7517984854 \\ 5, & \delta \in \{\delta_1^S, \delta_2^S\}, \quad \delta_2^S \approx +0.5159002918 \end{cases}$$

(ii) δ -Stabilität (numerisch)

getestet wurde nur im Bereich $[\delta_1^K, \delta_2^K]$, da außerhalb entweder keine Konvergenz vorliegt oder einer der Knoten negativ ist. Die Abbildungen zu diesem Beispiel, insbesondere Plot 11, zeigen eine noch schwächere δ -Abhängigkeit der Regionen $R_A(\delta)$ als in Beispiel 5.10.b, was die plausible Idee zur Minimierung von $|\sigma_2(C)|$ bestätigt.

1) Das zu $v(\delta)=-1$ in Beispiel 5.10.b Gesagte gilt hier entsprechend.

Die Situationen $\delta > 0$ und $\delta < 0$ sind verschieden, es gibt 2 verschiedene Grenzen zur A-Stabilität, die wiederum durch Intervallschachtelung numerisch bestimmt wurden, und zwar mit Plot 11,12,13 im Bereich $\delta > 0$, mit Plot 14,15,16 für $\delta < 0$. Die beiden Grenzen sind nur zufällig betragsmäßig in etwa gleich. Es ergibt sich der A-Stabilitätsbereich

$$-0.678 \leq \delta \leq +0.680 \quad .$$

Insbesondere liegt der Superkonvergenzparameter δ_2^S im Bereich der numerisch getesteten A-Stabilität.

5.13 Bemerkung

Abschließend sollen die Ergebnisse aus den Beispielen 5.10.b und 5.12 mit den in der Einleitung angesprochenen Fällen (5.10) verglichen werden, die in der Arbeit von Butcher [3] bzw. Wanner [44] untersucht sind. Bezüglich der Parameterkonfiguration sind die Beispiele paarweise gleich, die Einschränkung an die Lage der Kollokationsstellen ist jedoch verschieden. Sie "lautet" bei Butcher/Wanner singly-implicitness, führt in jedem Beispiel zu einer einparametrischen Schar von Verfahren und bedeutet, daß die bei der praktischen Durchführung, genauer bei der Lösung der zugeordneten impliziten Gleichungen, auftretende Matrix ein reelles Einpunktspektrum besitzt und daher besonders einfach "invertierbar" ist, da die Transformation auf Jordansche Normalform explizit angegeben werden kann.

Bei der Vorgehensweise hier, die sich nicht an Aufwandsfragen, sondern an Stabilitätseigenschaften orientiert, lautet die Einschränkung an die Zwischenknoten Schwerpunkt 1, was zu einer $(k-1)$ -parametrischen -im Fall $\#(K)=p-1$ und $d_i=1, i=1(1)p-1, (p-2)$ -parametrischen- Verfahrensschar führt (Diese ist in Beispiel 5.12 ($p=4$) auf Einparametrigkeit reduziert). Die folgende Tabelle diene zum Vergleich:

Paramterkonfiguration	$(0, t_1, t_2)$	$(0, t_1, t_2, t_3)$
<u>Butcher [3], Wanner [44]</u>		
Einschränkung	$t_1 = \lambda(3 - \sqrt{3})$, $t_2 = \lambda(3 + \sqrt{3})$, $\lambda > 0$	$t_1 = \lambda * 0.936...$, $t_2 = \lambda * 3.305...$, $t_3 = \lambda * 7.756...$, $\lambda > 0$
A-Stabilität (theoretisch)	$\lambda > 0.712..$, d.h.: $t_1 > 0.902..$, $t_2 > 3.368..$,	$\lambda > 0.892..$, d.h.: $t_1 > 0.834..$, $t_2 > 2.948..$, $t_3 > 6.918..$



Fuchs, hier, Beispiele 5.10.b, 5.12		$t_1 = 1 - \delta,$ $t_2 = 1 + \frac{1}{2}(1 - \frac{1}{\sqrt{2}})\delta,$ $t_3 = 1 + \frac{1}{2}(1 + \frac{1}{\sqrt{2}})\delta,$ $-1.171.. < \delta < 1$
Einschränkung	$t_1 = 1 - \delta,$ $t_2 = 1 + \delta, 0 < \delta < 1$	
A-Stabilität (theoretisch) falls	$0 < \delta \leq 0.130, \text{d.h.:}$ $1 > t_1 \geq 0.870,$ $1 < t_2 \leq 1.130$	-
A-Stabilität (numerisch)	$0 < \delta \leq 0.330, \text{d.h.:}$ $1 > t_1 \geq 0.670,$ $1 < t_2 \leq 1.330$	$-0.678 \leq \delta \leq +0.680, \text{d.h.:}$ $1.678 \geq t_1 \geq 0.320,$ $0.901 \leq t_2 \leq 1.100,$ $0.421 \leq t_3 \leq 1.580$

Insbesondere in Beispiel 5.12 ist also mit der hier vorgeschlagenen Kopplung der Kollokationsknoten ein erheblicher (eindimensionaler) Parameterbereich erzielt, für den A-Stabilität vorliegt. Man kann hier eigentlich nicht mehr von Konfluenz sprechen, sollte eher von "Fernwirkung" der Stabilitätsregion der (3,1)- α - β -approximation reden. Das Ziel, A-Stabilität mit in der Nähe und innerhalb des "aktuellen Integrationsintervalls" liegenden Zwischenknoten zu "halten", wurde erreicht. Die drei Beispiele 5.11.a,b und 5.12 sind als exemplarische Ausführung der im Zusammenhang zu K 5.9 und L 5.11 geäußerten Ideen zu sehen.

Bei den Ergebnissen von Butcher/Wanner fällt dagegen auf, daß die Knoten schon bei den obigen einfachen Beispielen zum Teil relativ weit außerhalb des aktuellen Intervalls liegen, wenn man A-Stabilität verlangt. Das gilt in noch viel stärkerem Maße für die komplizierteren, numerisch auf A-Stabilität getesteten Verfahren in der Arbeit [2] von Burrage/Chipman. Beispielsweise für $p=5, q=2$ (und $d_5=0, d_1=1, i=1(1)4$), erzwingt A-Stabilität

$$t_1 \geq 1.23, t_2 \geq 3.42, t_3 \geq 6.92, t_4 \geq 12.6.$$

Aufgrund der Bedeutung der Zwischenknoten (Information über die Ableitung der Lösung im aktuellen Integrationsbereich) ist zu befürchten, daß die Auflösung der zugeordneten impliziten Gleichungen Schwierigkeiten bereitet und vielleicht trotz A-Stabilität auch in der stationären Phase eines steifen Systems kleine Schrittweiten erforderlich sind. Sollte sich diese Vermutung bestätigen -praktische Tests liegen auch in [2] noch nicht vor-, so müßte man zu dem Schluß kommen, daß A-Stabilität eine zu harte Restriktion an ein Verfahren aus der von Butcher in [3] beschriebenen Teilklasse von $V^>$ ist.

5-3 A_{∞}^- , A_0^- und A -Stabilität

5.14 Definition

Sei $U^{[0]} \in V$. $U^{[0]}$ heie A_0 -stabil : $\iff R_A > R_- := \{z \in \mathbb{R} \mid z < 0\}$ ●

Die beiden Eigenschaften A_∞ -Stabilitt sowie A_0 -Stabilitt lassen sich sehr viel einfacher berprfen als A -Stabilitt, nmlich im wesentlichen durch Berechnung von Polynomnullstellen. Die Bedeutung der A_∞ -Stabilitt wurde zu Beginn von Kapitel 4 erwhnt. Die eine jener besagten Eigenschaften A_∞ -stabiler Verfahren, nmlich, da solche in vielen Fllen "fast A -stabil" oder sogar A -stabil sind, gilt fr A_0 -Stabilitt entsprechend.

Weiterhin ist es plausibel, da fr den Fall einer Verfahrensschar $(U^{[0]}(\delta))_{\delta \in \mathbb{R}}$ gem (5.1), (5.2), (5.13), fr die zu einem gegebenen Intervall $I \subset \mathbb{R}$ die Bedingungen

- (i) $g(A_2(\delta)) < 1$, $\delta \in I$,
- (ii) $R_A(\delta) \cap \mathbb{C}^-$ ist fr alle $\delta \in I$ konvexes, einfach zusammenhngendes Gebiet in $\mathbb{C}$,

gelten, die in $\text{int}(I)$ liegenden Grenzen zu A_∞ -, A_0 - und A -Stabilitt zusammenfallen. Es sei vermerkt, da (ii) zwar hufig auftritt, Kriterien dafr aber unbekannt sind. Die obige Plausibilittsbetrachtung soll mathematisch nicht genauer untersucht, jedoch durch Berechnung der Grenzen zur A_∞ - und A_0 -Stabilitt der Beispiele 5.10.a,b und 5.12 veranschaulicht werden.

Ziel dieses letzten Abschnittes ist es also, zu den bisherigen Angaben ber die Eigenschaften obiger Beispiele die Bereiche mit A_∞ - bzw. A_0 -Stabilitt hinzuzufgen und einige Grenzen zur A -Stabilitt (gem obiger Betrachtung) exakt anzugeben.

Dazu die beiden folgenden Aussagen:

5.15 Satz

Gegeben sei eine Schar $(U^{[0]}(\delta))_{\delta \in \mathbb{R}}$ mit $q=d_p+1$ gem (5.1), (5.2), (5.13). Man setze

$$\begin{aligned} \psi_\sigma(\delta) := & \prod_{j \in K} (c_j \delta)^{2d_j + \beta'^2} \prod_{j \in K} (1 - c_j \delta)^{2d_j + \sigma \beta'(2+r')} \prod_{j \in K} ((1 - c_j \delta) c_j \delta) \\ & + \sigma \beta' \sum_{j \in K} d_j ((1 - c_j \delta) c_j \delta)^{d_j - 1} \prod_{\substack{l \in K \\ l \neq j}} ((1 - c_l \delta) c_l \delta)^{d_l}, \end{aligned}$$

$$\bar{\psi}_{\sigma}(\delta) := \sigma \beta' \prod_{j \in K} (1 - c_j \delta)^{d_j} - \prod_{j \in K} (c_j \delta)^{d_j}, \quad \sigma = -1, 1, \text{ mit}$$

$$\beta' := \prod_{i \notin K} \left(\frac{t_i}{1-t_i} \right)^{d_i}, \quad \Gamma' := 2d_p + \sum_{i \notin K} \frac{d_i}{t_i(1-t_i)}.$$

Die Grenzen zur A_{∞} -Stabilität¹⁾ (bzgl. Variation von δ in $\mathbb{R} - \{c_j^{-1} | j \in K\}$) sind charakterisiert durch die Vereinigung der reellen, von Werten aus $\{c_j^{-1} | j \in K\}$ verschiedenen Nullstellen δ^* von $\psi_1, \bar{\psi}_1$, für die $|\beta(\delta^*)| \geq 1$ ist, mit denjenigen von $\bar{\psi}_1, \bar{\psi}_{-1}$, für die $\Gamma(\delta^*) \in (-4, 0)$ gilt.

Bemerkungen

- a) Im Spezialfall $K = \{j\}$ für ein $j \in \{1, \dots, p-1\}$ ergibt sich die Aussage aus S 4.7.a. Der Satz wurde in Kapitel 4 noch nicht in dieser Allgemeinheit formuliert, um nicht die Definition der Konfluenz vorziehen zu müssen.
- b) Der Aufwand zur Bestimmung der δ -Intervalle mit A_{∞} -Stabilität und derer mit Konvergenz besteht (im wesentlichen) aus der Berechnung der Nullstellen von 2 Polynomen aus $\prod_{2k}$ sowie 2 Polynomen aus $\prod_k$ ($k = d^{(K)}$).

Beweis

Analog zum Beweis von S 4.7.a.

Es ist bekannt, daß man bei Verfahren, denen bei Anwendung auf (T_{μ}) eine charakteristische Relation der Form (3.4) mit reellen Koeffizienten a_{ij} , $i=0(1)K$, $j=0(1)M$, zugeordnet ist, A_0 -Stabilität mittels des Routh- oder Hurwitz- Kriteriums (Vgl. Routh [36], Hurwitz [14] oder z.B. Obreschkoff [34]) überprüfen kann. Dazu ist auf (3.4) die in B 4.6 angesprochene Möbiustransformation anzuwenden. Die wurde von Friedli/Deltsch [6], Theorem 1, ausgeführt und wird jetzt, zugeschnitten auf die hiesigen Bezeichnungen im Fall $q=d_p+1$, formuliert:

¹⁾ Genauer: Diejenigen Werte δ^* mit $q(\psi(\infty, \delta^*))=1$ (Vgl. S 4.4).

5.16 Satz (i.wes.: Friedli/Jeltsch)

Gegeben sei $U^{[0]} \in V^>$ mit $q=d_p+1$, $g(A_2) < 1$ und $B \cap \mathbb{R}_- = \emptyset$. Ferner seien

$$(5.15) \quad \begin{cases} b_0(z) := (\beta-1)(q_0(z) + p_0(z)) - (q_1(z) + p_1(z)), \\ b_1(z) := 2(\beta q_0(z) + p_0(z) + p_1(z)), \\ b_2(z) := (\beta+1)(q_0(z) - p_0(z)) + (q_1(z) - p_1(z)), z \in \mathbb{C}. \end{cases}$$

Dann folgt die Äquivalenz: $U^{[0]}$ ist genau dann A_0 -stabil, wenn gilt:

$$(5.16) \quad \text{sign } b_j(z) = \text{sign } \beta, \quad j=0(1)2, \quad z \in \mathbb{R}_-.$$

Bemerkungen

a) Ohne die Voraussetzung $B \cap \mathbb{R}_- = \emptyset$ gilt nach wie vor, daß die Bedingungen in (5.16) A_0 -Stabilität von $U^{[0]}$ implizieren.

b) Sei

$$b_j(z) =: \sum_{l=0}^{m-q} \beta_{lj} z^l, \quad j=0(1)2. \quad \text{Dann gilt:}$$

(i) Hinreichend für (5.16) sind die beiden folgenden Bedingungen:

1. $\forall j \in \{0,1,2\} \exists l \in \{0, \dots, m-q\} : \beta_{lj} \neq 0$.
2. $\text{sign } \beta_{lj} = (-1)^l \text{sign } \beta$, falls $\beta_{lj} \neq 0, l=0(1)m-q, j=0(1)2$.

(ii) Notwendig für (5.16) ist:

Für alle $j \in \{0,1,2\}$ gilt $\text{sign } \beta_{lj} = (-1)^l \text{sign } \beta$ für diejenigen unter den Koeffizienten β_{lj} , $l=0(1)m-q$, die unter den nichtverschwindenden den höchsten oder niedrigsten Index l besitzen.

c) Im Fall einer Verfahrensschar $(U^{[0]}(\delta))_{\delta \in \mathbb{R}}$ mit $q=d_p+1$ gemäß (5.1), (5.2), (5.13) ergeben sich mittels Bemerkung b(i) zumindest Teile der Parameterbereiche, für die A_0 -Stabilität vorliegt, aus den Nullstellen der Koeffizientenfunktionen β_{lj} , $l=0(1)m-q$, $j=0(1)2$, wie an (einem der) untenstehenden Beispiele(n) demonstriert wird.

d) Für diese Teilbereiche besteht der Aufwand -vgl. (5.13)- (im wesentlichen) in der Berechnung der Nullstellen von $3(m-d_p)-1$ Polynomen aus π_{γ_k} sowie 2 Polynomen aus π_k . Letztere sind wieder $\bar{\psi}_1, \bar{\psi}_{-1}$, dienen zur Überprüfung der Voraussetzung $g(A_2(\delta)) < 1$ und liefern -vgl. Bemerkung b zu S 5.15- ebenfalls die δ -Bereiche mit Konvergenz.

Beweis

Mit Beispiel a zu S 6.3, D 5.14, (5.15) sowie den Ausführungen in Friedli/Jeltsch [6], S.405 und $q(A_2) < 1$ folgt die Äquivalenz:

$$U^{[0]} \text{ ist } A_0\text{-stabil} \iff b_j(z) \neq 0, z \in \mathbb{R}_-, j=0(1)2.$$

Die Behauptung des Satzes ergibt sich aus der Stetigkeit der b_j , $j=0(1)2$, und den Formeln (5.15) wegen $|b| = q(A_2)^{-1} > 1$ (Hierzu beachte man $b_2(z) = -z(\beta+1) + O(z^2)$, $z \rightarrow 0$).

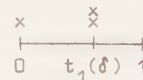
5.17 Beispiele

Im folgenden wird der Einfachheit halber in denjenigen Fällen mit $q(A_2) > 1$, auf die Satz 5 aus Mülthei [26] nicht anwendbar ist, $\gamma := -1$ "neu definiert" (Vgl. auch Bemerkung zum Fall $\gamma(\delta) = -1$ im Beispiel 5.10.b). Desgleichen bleibt die wegen S 6.4 unwesentliche Voraussetzung $B \cap \mathbb{R}_- = \emptyset$ aus S 5.16 unberücksichtigt.

a) (Vgl. Beispiele 4.8.b, 5.10.a, Plot 4.5.6)

$$\tilde{U} = (0, 1, 1), U^{[0]}(\delta) = (0, 1-\delta, 1-\delta), -\infty < \delta < 1$$

$$k=2, \sigma=1; \beta(\delta) = -(1-\delta)^2/\delta^2$$



(i) Konvergenz : Siehe Beispiel 5.10.a(i).

(ii) A_∞ -Stabilität: In Beispiel 4.8.b wurde bereits gezeigt:

$$U^{[0]}(\delta) \text{ ist } A_\infty\text{-stabil} \iff -\infty < \delta < \delta^\infty := \frac{1}{2}(1 - \frac{1}{\sqrt{3}}) \doteq \underline{0.2113248654}$$

(iii) A_0 -Stabilität : Für die Größen aus S 5.16 gilt:

$$p_0(z, \delta) = \frac{1}{\delta^2} [(-1+2\delta-2\delta^2) + \frac{2}{3}(1-\delta)(1-2\delta+\delta^2)] + \frac{z^2}{6} [-1+2\delta-\delta^2+4\delta^3-4\delta^4],$$

$$p_1(z, \delta) = \frac{2}{\delta^2} [(-1+2\delta) + \frac{2}{3}z(1-\delta+\delta^2)] + \frac{z^2}{6} [-1+2\delta(1-\delta+\delta^2)],$$

$$b_2(z, \delta) = \frac{1}{\delta^2} [(1-2\delta)z - \frac{z^2}{6} [1-6\delta+6\delta^2)].$$

Die beiden unterstrichenen Polynome ergeben sich ebenfalls bei der Berechnung der Grenzen zur A_∞ -Stabilität gemäß S 5.15.

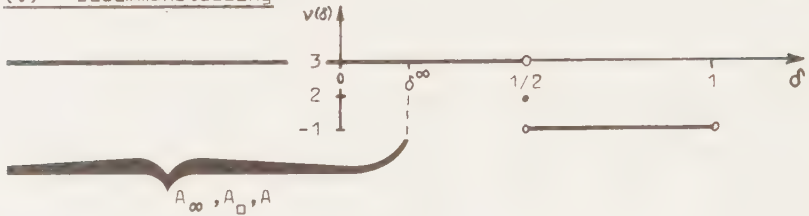
S 5.16 und Bemerkung b(i),(ii) dazu ergeben die Äquivalenz:

$$U^{[0]}(\delta) \text{ ist } A_0\text{-stabil} \iff -\infty < \delta \leq \delta^\infty = \frac{1}{2}(1 - \frac{1}{\sqrt{3}})$$

Die zum Beispiel gehörenden Zeichnungen Plot 4.5.6 legen die Vermutung nahe, daß die in der Einleitung zu diesem Abschnitt angesprochene Situation (i),(ii) hier für $I := (-\infty, \frac{1}{2})$ tatsächlich vorliegt. Offensichtlich hat man durch das Losen einer quadratischen Gleichung den Bereich der A -Stabilität bestimmt (Vgl.bsp. 4.1.a(ii)).

(iv) A-Stabilität : Siehe Beispiel 5.10.a(ii).

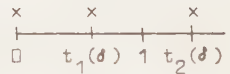
(v) Zusammenstellung



b) (Vgl. Beispiel 5.10.b, Plot 8.9, 10)

$$\tilde{U} = (0, 1, 1), U^{[0]}(\delta) = (0, 1-\delta, 1+\delta), |\delta| < 1$$

$$k = \sigma = 2; \quad \beta(\delta) = (1-\delta^2)/\delta^2$$



(i) Konvergenz : Siehe Beispiel 5.10.b(i).

(ii) A_∞-Stabilität: Aus S 5.15 und S 4.4 folgt:

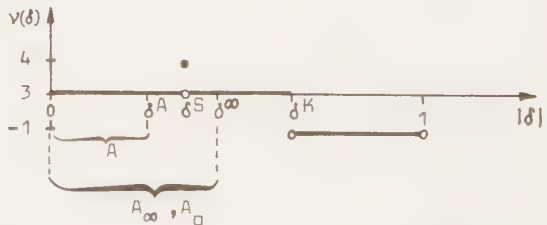
$$U^{[0]}(\delta) \text{ ist } A_\infty\text{-stabil} \iff |\delta| < \delta^\infty := \frac{1}{2}(\sqrt{3}-\sqrt{5}) \doteq 0.4370162445$$

(iii) A_0-Stabilität : S 5.16 und Bemerkung b(i), (ii) dazu ergeben auch hier die Äquivalenzaussage:

$$U^{[0]}(\delta) \text{ ist } A_0\text{-stabil} \iff |\delta| \leq \delta^0$$

(iv) A-Stabilität : Siehe Beispiel 5.10.b(ii).

(v) Zusammenstellung



$$\delta^K := 1/\sqrt{2} \doteq 0.7071067812, \delta^S := \sqrt{(5-\sqrt{13})/12} \doteq 0.3408871666, \delta^A := 0.330.$$

In diesem Beispiel ist also der Bereich der A_∞ - bzw. A_0 -Stabilität deutlich größer als derjenige, für den A-Stabilität vorliegt. Sogar für $0.331 \leq |\delta| < \delta^S$, wo der "Stabilitätsrand in einer Umgebung von Null nach rechts geöffnet" ist und A_0 -Stabilität vorliegt, hat man keine A-Stabilität, wie Plot 9.10 verdeutlichen. Situation (ii) aus der Einleitung dieses Abschnittes liegt offensichtlich in keinem Intervall, welches δ^A und δ^∞ im Innern enthält, vor.

c) (Vgl. Beispiel 5.12, Plot 11-16)

$$\tilde{U} = (0, 1, 1, 1), U^{[0]}(\delta) = (0, 1-\delta, 1+\frac{1}{2}(1-\frac{1}{\sqrt{2}})\delta, 1+\frac{1}{2}(1+\frac{1}{\sqrt{2}})\delta); \quad k=3, \sigma=2$$

Auch hier sollen nur solche δ interessieren, für die $t_j(\delta) > 0$, $j=1,2,3$, gilt. Laut Beispiel 5.12 ist dies genau für $\delta \in J$,

$$J := \left(-\frac{2\sqrt{2}}{1+\sqrt{2}}, 1 \right) \doteq (-1.172, 1),$$

der Fall. Weiter sollen zusätzlich zu den bereits gefundenen Eigenschaften nurmehr der Bereich der A_∞ -Stabilität ermittelt werden, welcher vermutlich mit dem Bereich der A_0 -Stabilität zusammenfällt. Mit S 5.15 ergibt sich

$$64 \psi_\sigma(\delta) = \delta^6(2-2\sigma) + 7\delta^5(2-2\sigma) + 49\delta^4(1+\sigma) + 8\delta^3(-2+5\sigma) - 56\delta^2(2+\sigma) + 64, \\ 8 \bar{\psi}_\sigma(\delta) = \delta^3(1-\sigma) - 7\sigma\delta^2 + 8\sigma, \quad \sigma = -1, 1$$

(Nullstellen von $\bar{\psi}_1, \bar{\psi}_{-1}$ vgl. Beispiel 5.12(i), δ_j^K , $j=1(1)5$).

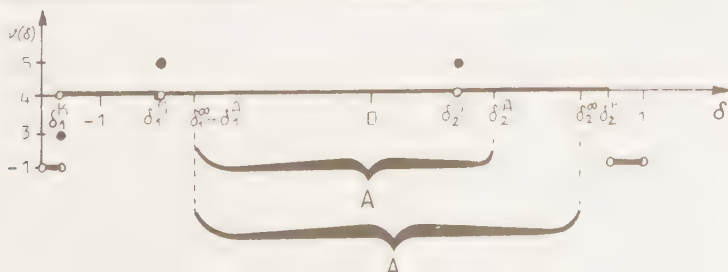
Zur Bestimmung der Nullstellen von ψ_1, ψ_{-1} (sowie derer von $\bar{\psi}_1, \bar{\psi}_{-1}$) wurde, wie gesagt, ein Programm von Werner [46] verwendet.

Es folgt -bei Beschränkung auf J - mittels S 5.15 die Äquivalenz:

$$U^{[0]}(\delta) \text{ ist } A_\infty\text{-stabil} \iff \delta \in (\delta_1^\infty, \delta_2^\infty), \\ \delta_1^\infty \doteq -0.6780103017, \quad \delta_2^\infty \doteq +0.8732311279.$$

Die Zeichnungen Plot 11, 12, 13 zeigen, daß in einem beliebigen Intervall I , welches im Innern sowohl die Grenze $\delta_2^A \doteq 0.680$ zur A -Stabilität als auch δ_2^∞ enthält, die Situation (ii) aus der Einleitung zu diesem Abschnitt offensichtlich nicht vorliegt (Entsprechend Beispiel b). Gagegen legen Plot 14, 15, 16 die Vermutung nahe, daß -entsprechend Beispiel a- in einer Umgebung von δ_1^∞ jenes (ii) zutrifft, womit $\delta_1^A \doteq -0.678$ und δ_1^∞ zusammenfielen.

Offensichtlich ist die eine Grenze zur A -Stabilität, δ_1^A , Nullstelle des Polynoms $\psi_1 \in \Pi_4$. Zusammenstellung:



(Bzgl. $\delta_1^K, \delta_2^K, \delta_1^S, \delta_2^S$ siehe Beispiel 5.12, $\delta_1^\infty, \delta_2^\infty, \delta_1^A, \delta_2^A$ siehe oben.)

5.18 Abschließende Bemerkungen

- a) Die vorstehenden Untersuchungen zur A_∞ - und A_0 -Stabilität sind in ähnlicher Weise für den allgemeinen Fall $q > d_p$ durchführbar. Dabei erhöhen sich nur Anzahl und Grad der zu diskutierenden Polynome (Vgl. auch B 4.6).
 - b) Für die praktische Anwendung sind im allgemeinen die Verfahren mit $d_i=1, i=1(1)p-1$, denen mit Vielfachheiten größer als 1 vorzuziehen, da sonst die rechte Seite f der Differentialgleichung (1.1) differenziert werden muß. In den Verfahrensscharen gemäß (5.1), (5.2) mit $d_p > 0$ oder $\#(K) > 1$ darf $\{\delta\}$ nicht zu klein gewählt werden, da sonst die $t_j(\delta), j \in K$, (und $t_p=1$) zu nahe zusammenliegen. Im Beispiel 5.17.c ist $\delta=\delta_2^S$ vorzuziehen, aber auch bei Anwendung des Verfahrens aus Beispiel 5.17.b sollte man den Superkonvergenzparameter wählen, wenn dieser auch leicht außerhalb des Bereichs der A-Stabilität liegt.
 - c) Im vorstehenden Abschnitt wurde gezeigt, wie die Bereiche der A_∞ - bzw. A_0 -Stabilität mit einfachen algebraischen Methoden, nämlich im wesentlichen durch Bestimmung von Polynomnullstellen, ermittelt werden können. Dabei fallen die gefundenen Intervallgrenzen zum Teil offensichtlich mit denjenigen für A-Stabilität zusammen. Es ist auch daher, daß man im allgemeinen mit "fast A-Stabilität" auskommt, für die Untersuchungen weiterer Beispiele durchaus sinnvoll, von der langwierigen, auf T 5.8 basierenden Abschätzungsmethode abzusehen. Statt dessen sollte man sich mittels der Untersuchungen der Konvergenzgebiete (Vgl. Beispiele 3.6), der Gleichungen aus S 5.15 sowie Computerzeichnungen der Stabilitätsregionen einen Überblick über die Eigenschaften der zu untersuchenden Verfahrensschar verschaffen, wie oben exemplarisch demonstriert.
- Dabei liefert T 5.8 mit K 5.9 eben die Anregung, wie eine Schar zu wählen ist, um die gewünschten Eigenschaften zu erzielen.

6.1 Lemma

Seien $M \in \mathbb{N}$, $w = (w_1, \dots, w_M) \in \mathbb{C}^M$, $y := (\sigma_1(w)), 1=0(1)M$,
 $y^* := (\sigma_1(w^*)), 1=0(1)M$, und

$$(6.1) \quad \begin{cases} \Omega^{(M)} := \left((-1)^j \binom{M-j}{1-j} \right), 1, j=0(1)M \in L(\mathbb{R}^{M+1}), \\ N(t) := \prod_{i=1}^M (t-w_i), t \in \mathbb{R}. \end{cases}$$

Dann gelten die folgenden Aussagen:

a) Für $w = (w_1, w_2)$ mit $w_1 \in \mathbb{C}^r$, $w_2 \in \mathbb{C}^{M-r}$, $1 \leq r \leq M-1$ ist

$$\sigma_1(w) = \sum_{j=0}^r \sigma_j(w_1) \cdot \sigma_{1-j}(w_2), 1 \in \mathbb{Z}.$$

b) $\sigma_1(\underbrace{(w, \dots, w)}_M) = \binom{M}{1} w^1, 1 \in \mathbb{Z}$, mit $\binom{M}{1} := 0, 1 \notin \{0, \dots, M\}$.

c) $\sigma_1(w^*) = \sum_{j=0}^1 (-1)^j \binom{M-j}{1-j} \sigma_j(w), 1=0(1)M$.

d) $(\Omega^{(M)})^2 = I, y = \Omega^{(M)} y^*.$

e) $N^{(j)}(0) = j! (-1)^{M-j} \sigma_{M-j}(w), N^{(j)}(1) = j! \sigma_{M-j}(w^*), j=0(1)M.$

f) Für die in (1.17) definierten Funktionen $\tilde{p}, \tilde{q}$ gelten die Darstellungen

$$\tilde{p}(z, w) = \sum_{l=0}^M \frac{(M-1)!}{M!} \sigma_1(w^*) z^l, \tilde{q}(z, w) = \sum_{l=0}^M \frac{(M-1)!}{M!} \sigma_1(w) (-z)^l.$$

Beweis

Elementar.

6.2 Lemma

Seien $k \in \mathbb{N}$, $\alpha \in \mathbb{C}$, $b, c \in \mathbb{C}^k$, $D \in GL(\mathbb{C}^k)$. Dann gilt:

$$\det \begin{pmatrix} \alpha & b^T \\ c & D \end{pmatrix} = \det(D) (\alpha - b^T D^{-1} c)$$

Beweis

Householder [13], S.17, (17).

6.3 Satz

Sei $U^{(0)} \in \mathcal{U}^>$.

a) Für die Matrix $\Psi(z)$ aus T 2.2 gilt für $z \in \mathbb{C} \setminus N_{Q_0}$ die Darstellung

$$(6.2) \quad \Psi(z) = \begin{pmatrix} \frac{p_0(z) - q_1(z)^T G^{-1} G^* \omega}{q_0(z)} & \frac{p_r(z) - q_r(z)^T G^{-1} G^* \Omega_r}{q_0(z)} \\ G^{-1} G^* \omega & G^{-1} G^* \Omega_r \end{pmatrix}$$

Hierbei sind $G, G^*, \Omega_r \in GL(\mathbb{R}^{q-d_p}), \omega \in \mathbb{R}^{q-d_p}$ definiert vermöge

$$(6.3) \quad \left\{ \begin{array}{l} G := \begin{pmatrix} 0 & \dots & \varepsilon \gamma_0 \\ 0 & \dots & -\varepsilon \gamma_0 & -\varepsilon \gamma_{-1} \\ \vdots & \vdots & \vdots & \vdots \\ (-1)^{q-d_p-1} \varepsilon \gamma_0 & \dots & \dots \end{pmatrix}, G^* := \begin{pmatrix} 0 & \dots & \gamma_0^* \\ 0 & \dots & \gamma_0^* & \gamma_{-1}^* \\ \vdots & \vdots & \vdots & \vdots \\ \gamma_0^* & \gamma_{-1}^* & \dots & \gamma_{1-d_p}^* \end{pmatrix}, \\ \varepsilon := (-1)^{m-d_p}, \gamma_{-i} := \sigma_{m-q-i}(\tau, \underbrace{1, \dots, 1}_{d_p}), \\ \gamma_{-i}^* := \sigma_{m-q-i}(\underbrace{1, \dots, 1}_{d_p}, \tau^*), \quad i \in \mathbb{N}_0, \\ \begin{pmatrix} 1 & 0 \\ \omega & \Omega_r \end{pmatrix} := \Omega := \underbrace{\Omega^{(q-d_p)}}_{(6.1)}. \end{array} \right.$$

b) Es gelten die folgenden Relationen:

$$(i) \quad q(z)^T \Psi(z) = p(z)^T, \quad z \in \mathbb{C} \setminus N_{Q_0},$$

$$(ii) \quad \det \Psi(z) = (-\beta)^{d_p-q} p_0^C(z) / q_0(z), \quad z \in \mathbb{C} \setminus N_{Q_0}, \text{ mit}$$

$$(6.4) \quad \beta := (-1)^{m-d_p} \prod_{i=1}^{p-1} \left(\frac{t_i}{1-t_i} \right)^{d_i},$$

$$p_0^C(z) := \tilde{p}(z, (\underbrace{1, \dots, 1}_{d_p}, \tau, \underbrace{1, \dots, 1}_q, m)), \quad z \in \mathbb{C},$$

$$(iii) \quad \text{rang } \Psi(z) \geq q-d_p, \quad z \in \mathbb{C} \setminus N_{Q_0}.$$

Beweis

- a) Aus (2.3) und L 6.1.a, e folgen die Identitäten $H = DG$ und $(h^*, h^*) = (DG^* \omega, DG^* \omega_r)$ mit $D := \text{diag}(d_p!, \dots, (q-1)!)$. Aus (2.2) folgt unmittelbar durch Invertierung der dort auftretenden "Fast-Elementarmatrix" die Darstellung

$$(*) \quad \Psi(z) = \begin{pmatrix} (p(z)^T - q_r(z)^T H^{-1}(h^*, h^*)) / q_0(z) \\ H^{-1}(h^*, h^*) \end{pmatrix}, \quad z \in \mathbb{C}^{N_{q_0}}.$$

Somit folgt (6.2) mit den Bezeichnungen aus (6.3).

- b) (i) : Trivial mit (*) und $q_l = (q_0, q_r^T)^T$.

(ii) : $\det G^{-1} G^* \omega_r = (-\beta)^{d_p - q}$ (Vgl. Beweis zu S 6.5.d). Mit L 6.2 folgt sofort

$$\det \Psi(z) = (-\beta)^{d_p - q} \frac{1}{q_0(z)} (p_0(z) - p_r(z)^T \omega_r^{-1} \omega).$$

L 6.1.f, (2.2) und L 6.1.a ergeben $p_0^C(z) = p(z)^T \omega_0$, $z \in \mathbb{C}$.

Aus L 6.1.d folgen $\omega_0 = e_0^* = (1, \omega^T)^T$ und $\omega_r^{-1} \omega = -\omega$, was die Behauptung ergibt.

(iii) : Trivial.

Beispiele

Mit β aus (6.4) sowie

$$(6.5) \quad \Gamma := 2d_p + \sum_{i=1}^{p-1} \frac{d_i}{t_i(1-t_i)}$$

gilt im Fall

- a) $q = d_p + 1$

$$\Psi = \frac{1}{\beta} \begin{pmatrix} \frac{\beta p_0 - q_1}{q_0} & \frac{\beta p_1 + q_1}{q_0} \\ 1 & -1 \end{pmatrix}, \quad f_\Psi(\lambda) = \lambda^2 - (R_0 - \frac{q_0^C}{\beta q_0}) \lambda - \frac{p_0^C}{\beta q_0},$$

$$q_0^C := \underbrace{\tilde{q}(\cdot, (0, \dots, 0, 1, 1, \dots, 1), m)}_{d_p} \underbrace{\quad}_{q},$$

- b) $q = d_p + 2$

$$\Psi = \frac{1}{\beta} \begin{pmatrix} \frac{\beta p_0 + (2+\Gamma)q_1 - q_2}{q_0} & \frac{\beta p_1 - (1+\Gamma)q_1 + q_2}{q_0} & \frac{\beta p_2 + \Gamma q_1 - q_2}{q_0} \\ -(2+\Gamma) & 1+\Gamma & -\Gamma \\ 1 & -1 & 1 \end{pmatrix},$$

$$f_{\psi}(\lambda) = \lambda^3 - \frac{\lambda^2}{\beta q_0} (\beta |p_0 + (2+r)(q_0 + q_1) - q_2) \\ + \frac{\lambda}{\beta^2 q_0} (\beta (2+r)(|p_0 + |p_1) - |p_2 + q_0^c) - \frac{\lambda^0}{\beta^3 q_0} \beta p_0^c.$$

(Der Einfachheit halber ist das Argument z weggelassen.)

6.4 Satz

Im Fall $\underline{q} = \underline{q}_0 + 1$ setze man mit den Bezeichnungen aus K 2.5. in $\mathbb{C} - N_{q_0}$:

$$\Delta := (a-d)^2 - 4ac,$$

$$M := \begin{cases} \frac{(a+d)^2}{a(d+c)}, & a+d \neq 0 \text{ oder } d+c \neq 0, \\ 4, & a+d = d+c = 0. \end{cases}$$

Dann gelten für die in D 2.6 und D 2.7 definierten Mengen die folgenden Charakterisierungen:

$$A = A' = N_{\mathbb{C} - N_{q_0}^c - q_0 p_0^c}, \quad A' = \{z \in \mathbb{C} - N_{q_0} \mid M(z) \in \{0, 4\}\},$$

$$B = \{z \in \mathbb{C} - N_{q_0} \mid p_0(z) q_0^c(z) - q_0(z) p_0^c(z) = 0, |\beta p_0(z)| < |q_0^c(z)|\},$$

$$C = \{z \in \mathbb{C} - N_{q_0} - N_{p_0^c - A' - B} \mid c(z) = 0, |a(z)| > |d(z)|\}$$

(Bzgl. β und p_0^c siehe S 6.3.b, bzgl. q_0^c siehe Beispiel a zu S 6.3).

Weiter gilt für $m \geq 2$:

$$\# \Delta, \# A', \# B, \# C, \# \dots < \infty.$$

Bemerkungen

a) Zum allgemeinen Fall $q > d_p$:

Bei A' handelt es sich im allgemeinen um (eine) eindimensionale komplexe Mannigfaltigkeit(en).

Sei zu $z \in \mathbb{C} - N_{q_0}$ $x_{q_0}^F(z)$ ein Eigenvektor zu $\lambda_0(z)$, erweitert zu einem System $\{x_j^F(z) \mid j=0(1)q-d_p\}$ linear unabhängiger Eigen- und Hauptvektoren der Frobeniusform von ψ und

$$\sum_{j=0}^{q-d_p} c_j(z) x_j^F(z) = (s_0(z), \dots, s_{q-d_p}(z))^T$$

die eindeutig bestimmte Zerlegung des Startvektors der Iteration

(I) (Siehe Einleitung zu Abschnitt 2-2). Es gilt

$$B \subset (\mathbb{C} - N_{q|_0}) \cap N_{c_0}^{(1)} \subset (\mathbb{C} - N_{q|_0}) \cap N_{c_0} \cdot \dots \cdot c_{q-d_p}$$

Zu $z \in \mathbb{C}$ ist gemäß Bemerkung a zu D 2.7 der Eigenvektor $y_0(z)$ der Matrix $V(z)$ zum Eigenwert $\lambda_0(z)$ bis auf Normierung eindeutig bestimmt. Mit (2.10) und dem Beweis zu T 2.8 folgt

$$C = (\mathbb{C} - N_{q|_0} - N_{c_0} A' - B) \cap N_{e_0}^T y_0$$

- b) Für den Beweis der für $q=d_p+1$ im obigen Satz konkret angegebenen Charakterisierungen müssen die Überlegungen unter a genauer ausgeführt werden, was im einzelnen zwar elementar, insgesamt jedoch langwierig ist und nicht näher aufgeschrieben werden soll.

6.5 Satz

Sei $U^{[0]} \in V^>$. Man betrachte die für die Konvergenz von $U^{[0]}$ "verantwortliche" Matrix $A_2 \in L(\mathbb{R}^{q-d_p})$ (Mülthei [28], S.452 und Satz 4) sowie die Hermiteschen Grundpolynome $l_{0,j-1}$, $j=d_p+1(1)q$ ([28], S. 452). Es sei $U_p^{[0]} := (0, U^{[0]}) \in V^>$, die zugeordneten Größen werden mit einem (ggf. weiteren) Index p versehen.

Dann gelten mit den Bezeichnungen aus T 2.2 und S 6.3 die folgenden Beziehungen und Darstellungen:

$$a) \quad (G_p^{-1} G_p^* \Omega_{r,p})_r = G^{-1} G^* \Omega_r ; \quad (H_p^{-1} H_p^*)_r = H^{-1} H^* ,$$

$$b) \quad (l_{0,d_p}, \dots, l_{0,q-1}) = v_r^T G^{-1} D^{-1} = v_r^T H^{-1} ,$$

$$c) \quad A_2 = D G^* \Omega_r^T G^{-1} D^{-1} = H^* H^{-1}, \quad D := \text{diag}(d_p!, \dots, (q-1)!)$$

$$(\text{Insbesondere ist } A_2 \sim \psi_r \text{ und } A_2 \in GL(\mathbb{R}^{q-d_p})) ,$$

$$d) \quad \det A_2 = (-\beta)^{d_p - q} .$$

Korollar

- a) Bei Symmetrie der inneren Kollokationsstellen einschließlich ihrer Vielfachheiten ($T=T^*$) gilt $Q(A_2) \gg 1$ (Eine ähnliche

1) Wohldefiniert ist -genauer- nur $(\mathbb{C} - N_{q|_0} - A') \cap N_{c_0}$, was jedoch für die Anwendung im Beweis zu L 3.3.b genügt.

Aussage beinhaltet Korollar 1 zu Satz 8 in Mülthei [26]).

- c) Betrachtet man den Grenzübergang von α der insgesamt $m-q-d_p$ inneren Kollokationsstellen (einschließlich ihrer Vielfachheiten) gegen 1, wobei genauer die Differenz dieser Zwischenknoten zu 1 $O(\delta), \delta \rightarrow 0$, sei¹⁾, so gilt die Abschätzung

$$g(A_2(\delta)) = O(\delta^{m-(q+d_p-1)}, \delta \rightarrow 0.$$

Offensichtlich ist es also zur Erzielung von Konvergenz günstig, Kollokationsstellen in die Nähe von 1 zu legen (Vgl. hierzu auch Mülthei [26], Hilfssatz 7 und Satz 7).

Beispiele

Mit β aus (6.4), Γ aus (6.5) und $\Gamma' := \frac{\gamma-1}{\gamma_0} \Gamma + \frac{\gamma^*}{\gamma_0^*} - \frac{\gamma-2}{\gamma_0}$ (vgl. (6.3)) gilt für

a) $q = d_p + 1$

$$H^{-1}H^* = \frac{1}{\beta} (-1),$$

b) $q = d_p + 2$

$$H^{-1}H^* = \frac{1}{\beta} \begin{pmatrix} 1+\Gamma & -\Gamma \\ -1 & 1 \end{pmatrix},$$

c) $q = d_p + 3$

$$H^{-1}H^* = \frac{1}{\beta} \begin{pmatrix} -(1+2\Gamma+\Gamma') & \Gamma+\Gamma' & -\Gamma' \\ 2+\Gamma & -(1+\Gamma) & \Gamma \\ -1 & 1 & -1 \end{pmatrix}.$$

Bzgl. der Aussage in S 6.5.a beachte man $\Gamma_p = \Gamma$, $\beta_p = -\beta$.

Beweis des Korollars

- a) $T=T^*$, S 6.5.d und (6.4) implizieren $|\det A_2| = 1$, woraus $g(A_2) \geq 1$ folgt.
 b) Man überlegt sich zunächst mit (6.3), daß die Elemente von $G^{-1} = G^{-1}(\delta)$ für $\delta \rightarrow 0$ beschränkt sind. Ebenso aus (6.3) folgt

$$\gamma_{-i}^* = \gamma_{-i}^*(\delta) = O(\delta^{m-1}), \delta \rightarrow 0, i=0(1)d_p-1,$$

und damit $\|G^*\|_\infty = \|G^*(\delta)\|_\infty = O(\delta^{m-(q+d_p-1)}, \delta \rightarrow 0$. Die Aussage von S 6.5.c ergibt $g(A_2) = g(G^{-1}G^* \Omega_r) \leq \|G^{-1}G^* \Omega_r\|_\infty$.

¹⁾ Präzise Definition siehe (5.1) mit $d^{(K)} := \alpha$.

Beweis des Satzes

- a) Aus $\Omega_p = \Omega^{(q+1-d_p)}$ ergibt sich mittels (6.1) $\Omega_{r,p} = -\Omega$. Wegen der Dreiecksstruktur der drei Matrizen G, G^*, Ω_r folgt mit den Rechenregeln über Matrixinversion und -multiplikation unter Beachtung von (6.3) und $\epsilon_p = -\epsilon$ zunächst $(G_p^{-1} G_p^*)_r = -G^{-1} G^*$ und damit $(G_p^{-1} G_p^* \Omega_{r,p})_r = G^{-1} G^* \Omega_r$. Der zweite Teil der Aussage ist entweder analog oder über den Zusammenhang zwischen H, H^* und G, G^* (Siehe Beweis zu S 6.3.a) einzusehen.
- b) Sei mit (3.11) $h(t) := \tilde{N}(t)t^{-d_p}$, $t \in \mathbb{R}$. Die Definition der hermiteschen Grundpolynome in [28], S.452, besagt

$$1_{0,j-1}(t) = \frac{\tilde{N}(t)}{(j-1)!} \sum_{\alpha=0}^{q-j} t^{\alpha+j-d_p-1} \frac{1}{\alpha!} \left(D_t^\alpha \frac{1}{h(t)} \Big|_{t=0} \right) =$$

$$\frac{v_r(t)^T}{(j-1)!} \left(\underbrace{\frac{-1}{(q-j)!} \left(D_t^{q-j} \frac{1}{h(t)} \Big|_{t=0} \right)}_1, \dots, \underbrace{\frac{(-1)^{q-j+1}}{0!} \left(D_t^0 \frac{1}{h(t)} \Big|_{t=0} \right)}_{q-j+1}, 0, \dots, 0 \right) \quad \nearrow \quad q-d_p$$

$j=d_p+1(1)q$. Wie man sich leicht überlegt, bleibt

$$G^{-1} = \left((-1)^i \left(\frac{D_t^{q-d_p+1-i-j}}{(q-d_p+1-i-j)!} \frac{1}{h(t)} \Big|_{t=0} \right) \right), i, j=1(1)q-d_p$$

($z! := \infty$, $-z \in \mathbb{N}$), zu zeigen, was man durch verifizieren der Relation $G^{-1}G = I_r$ nachrechnet (Man beachte weiter $H = DG$).

- c) Aus der Definition von A_2 (Mülthei, [28], S.452) folgt:

$$A_2 = \left(\frac{d^{d_p}}{dt^{d_p}}, \dots, \frac{d^{q-1}}{dt^{q-1}} \right)^T (1_{0,d_p}(t), \dots, 1_{0,q-1}(t)) \Big|_{t=1} = b$$

$$\left(\frac{d^{d_p}}{dt^{d_p}}, \dots, \frac{d^{q-1}}{dt^{q-1}} \right)^T v_r(t)^T \Big|_{t=1} \stackrel{H^{-1}}{=} H^* H^{-1} \quad (2.3)$$

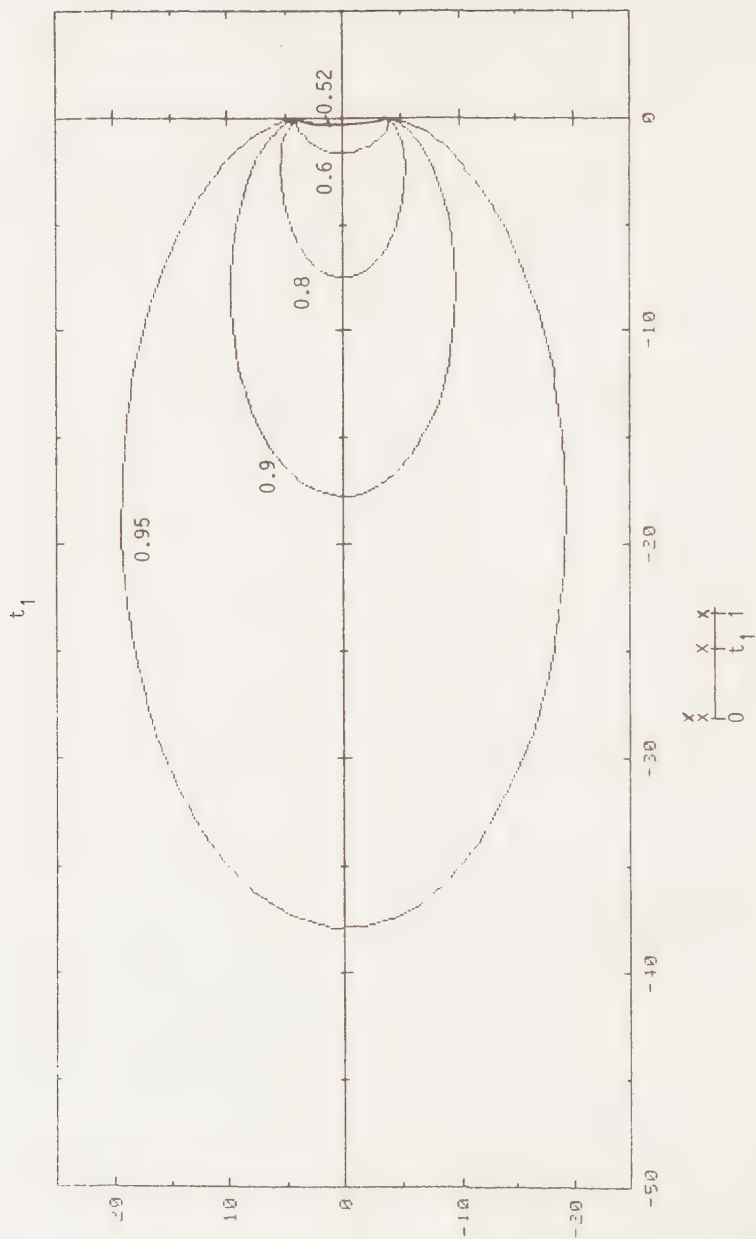
Für die Zerlegung $A_2 = DG^* \Omega_r^{-1} G^{-1} D^{-1}$ beachte man $H^* = DG^* \Omega_r$, $H = DG$ (Siehe Beweis zu S 6.3.a) sowie $\Omega_r^2 = I_r$.

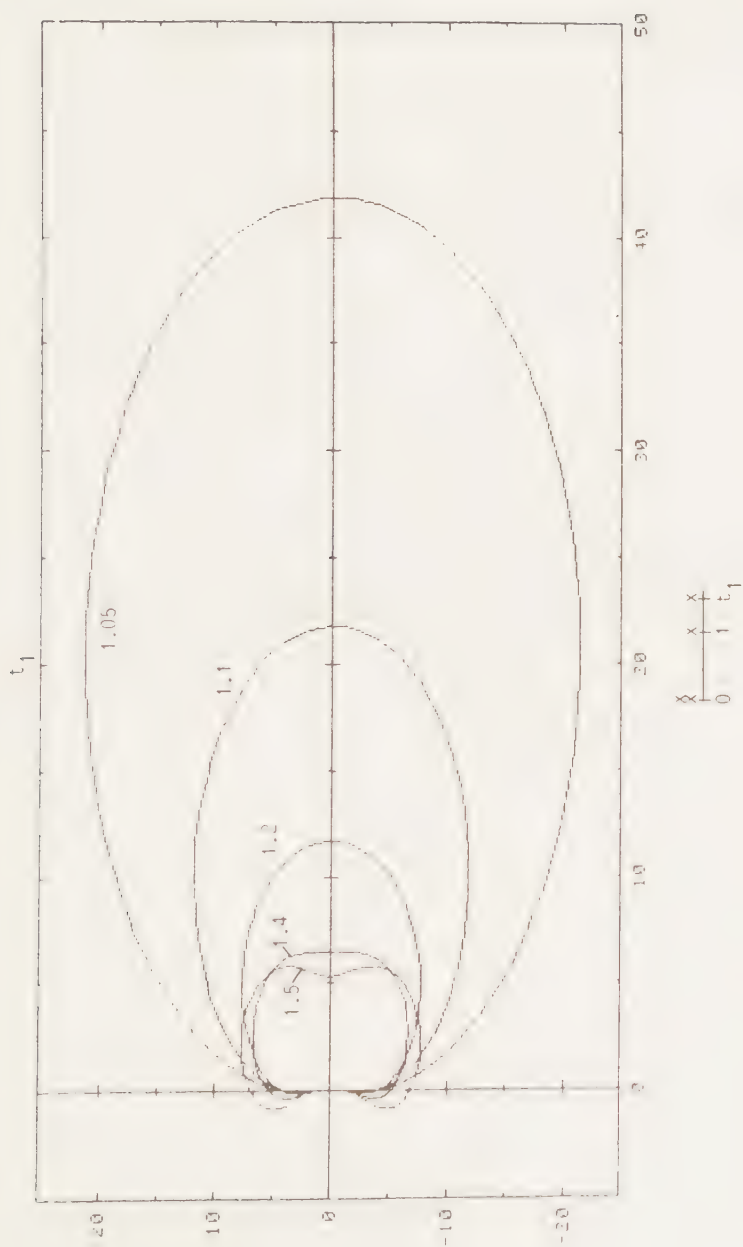
- d) Die unter c bewiesene Aussage impliziert

$$\det A_2 = (\det G)^{-1} \det G^* \det \Omega_r ;$$

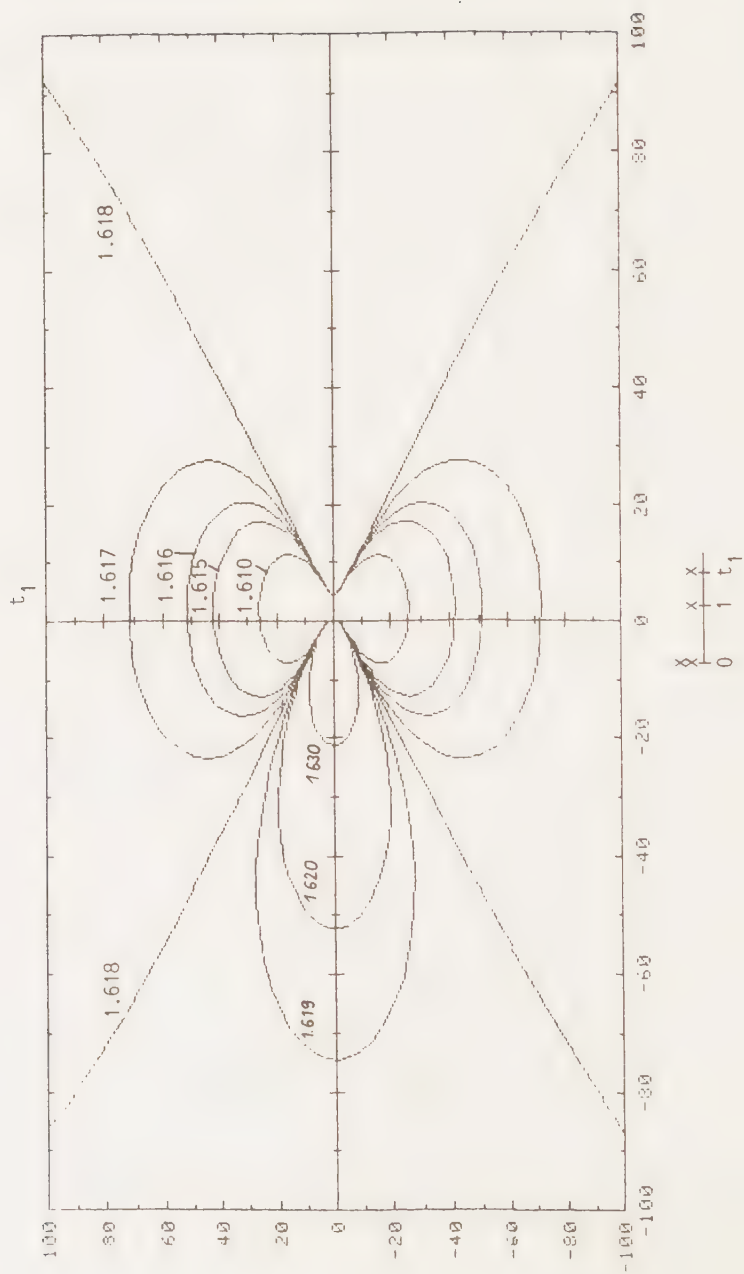
die Behauptung folgt aus den Definitionen der Matrizen in (6.3) und (6.1).

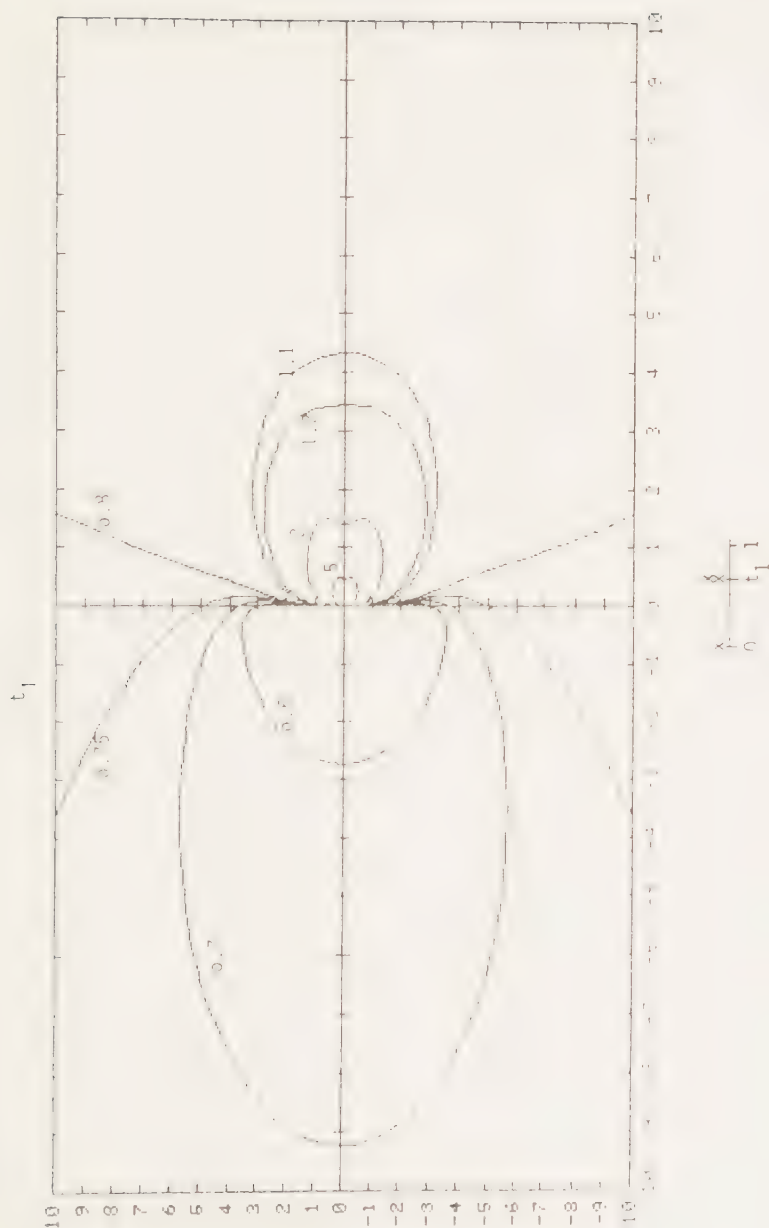
Abbildungen



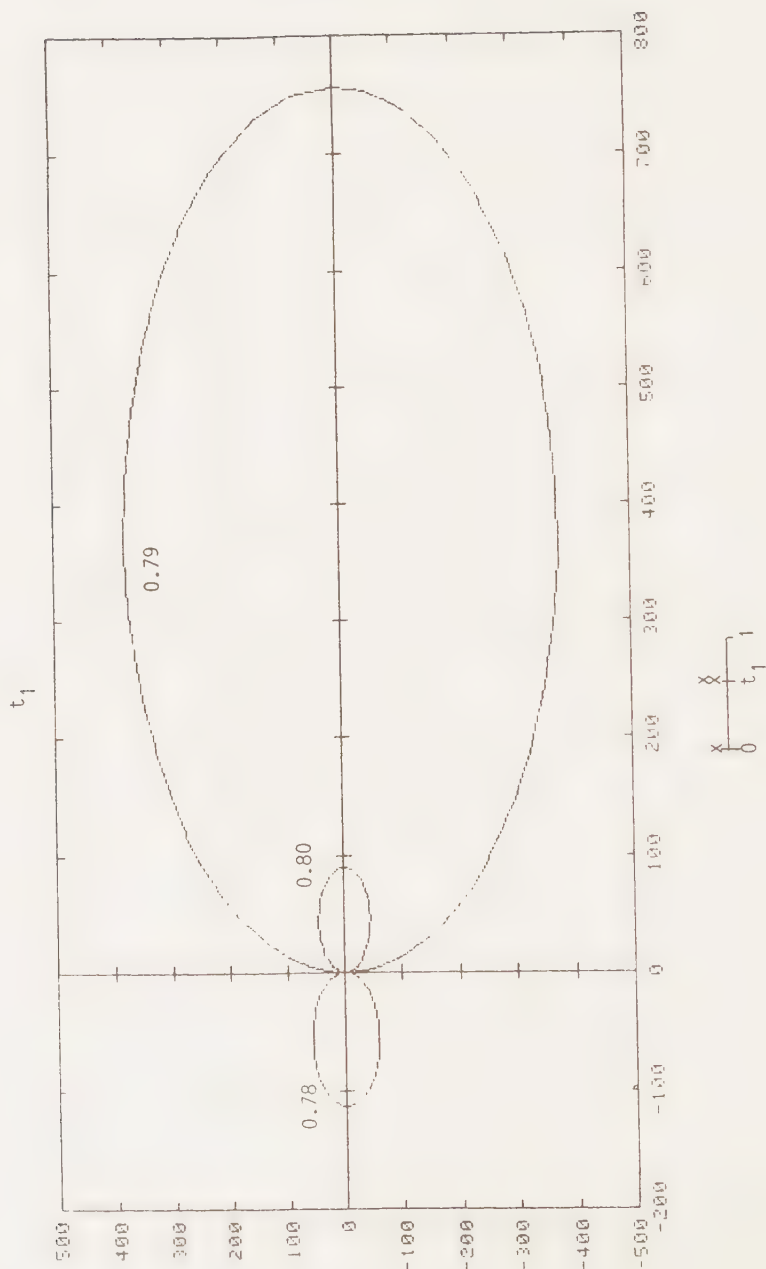


Plot 3

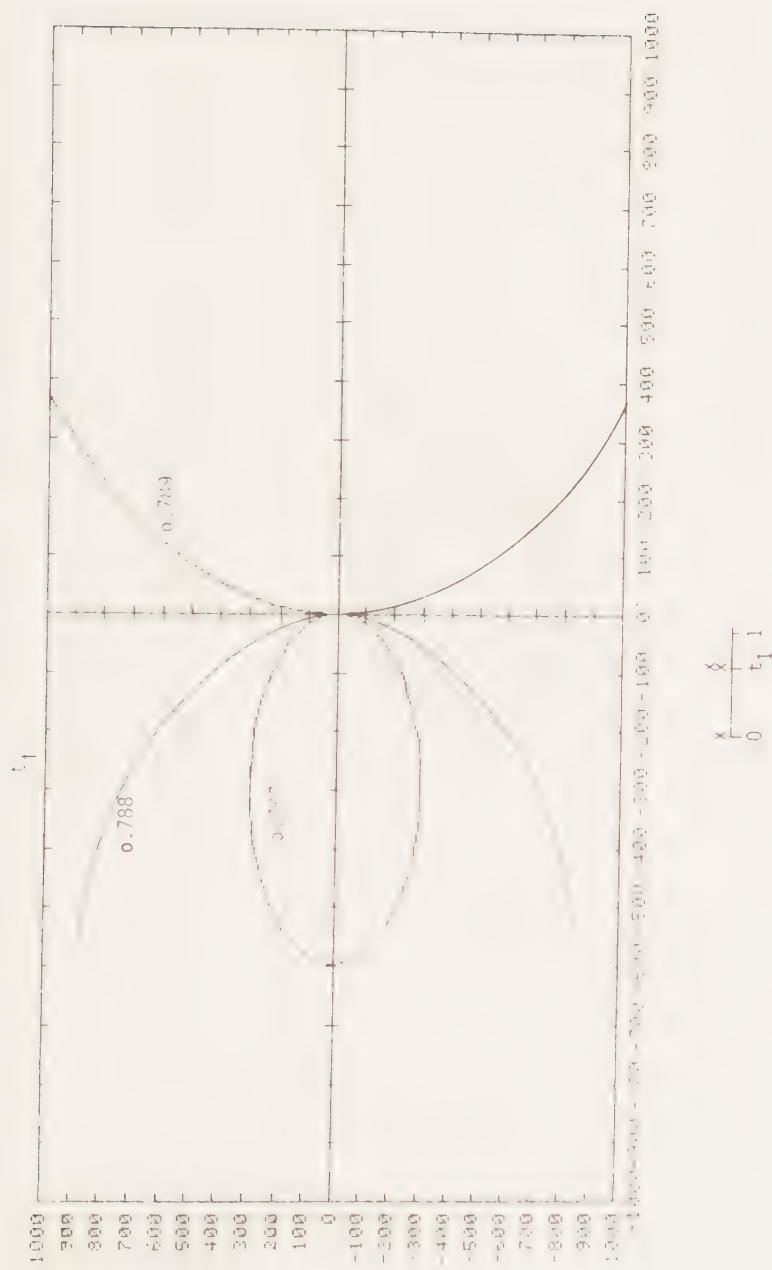




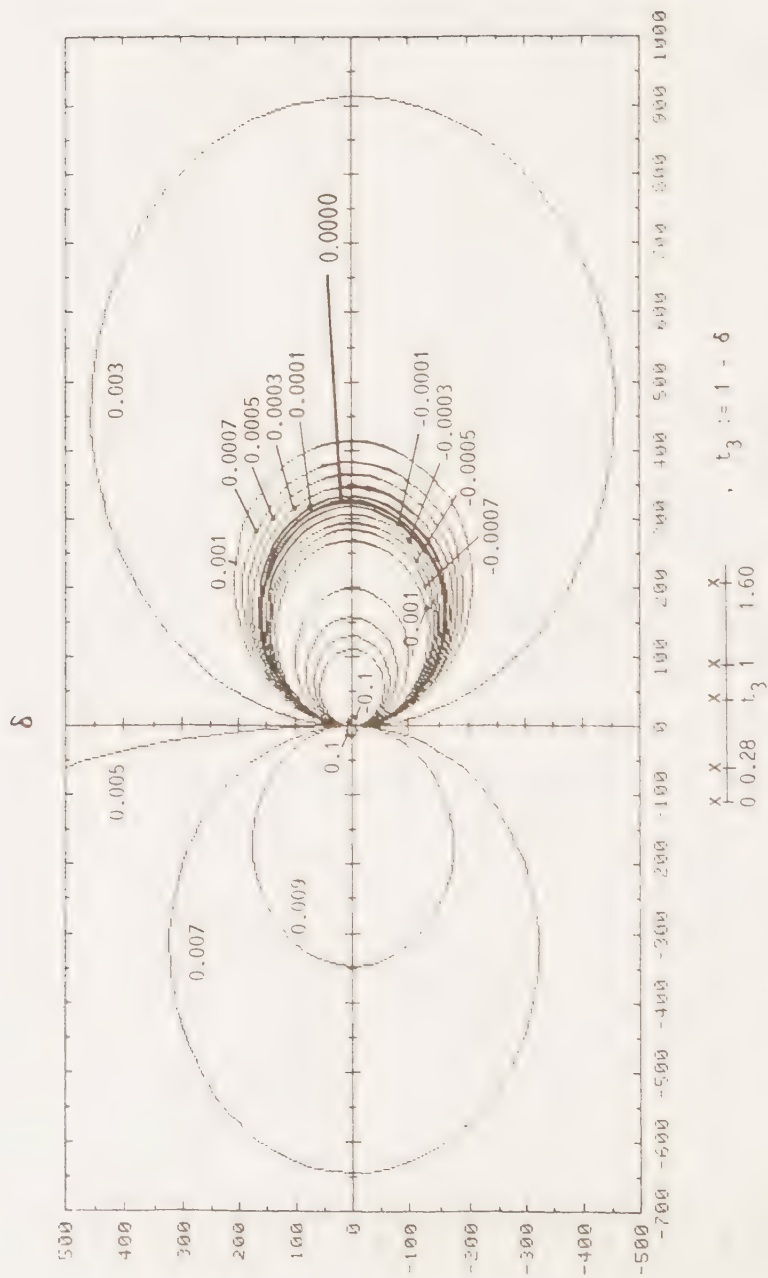
Plot 5



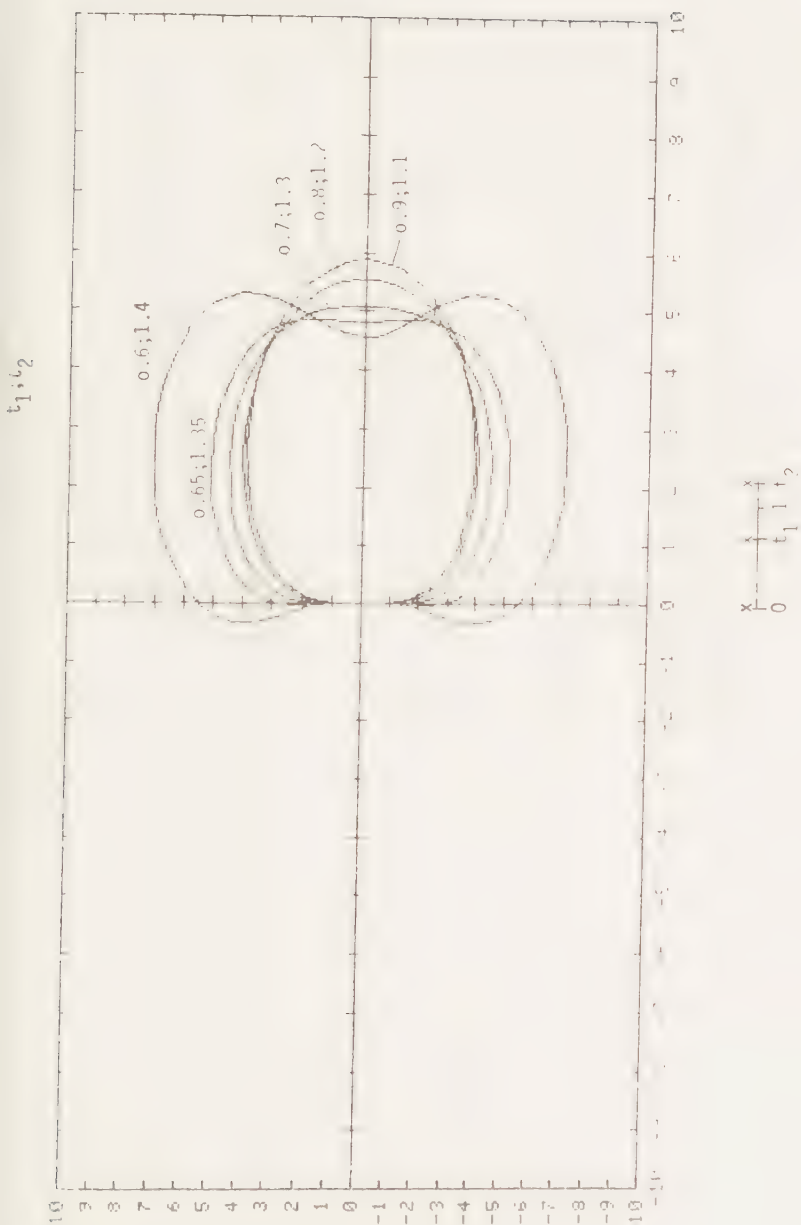
Plot 6



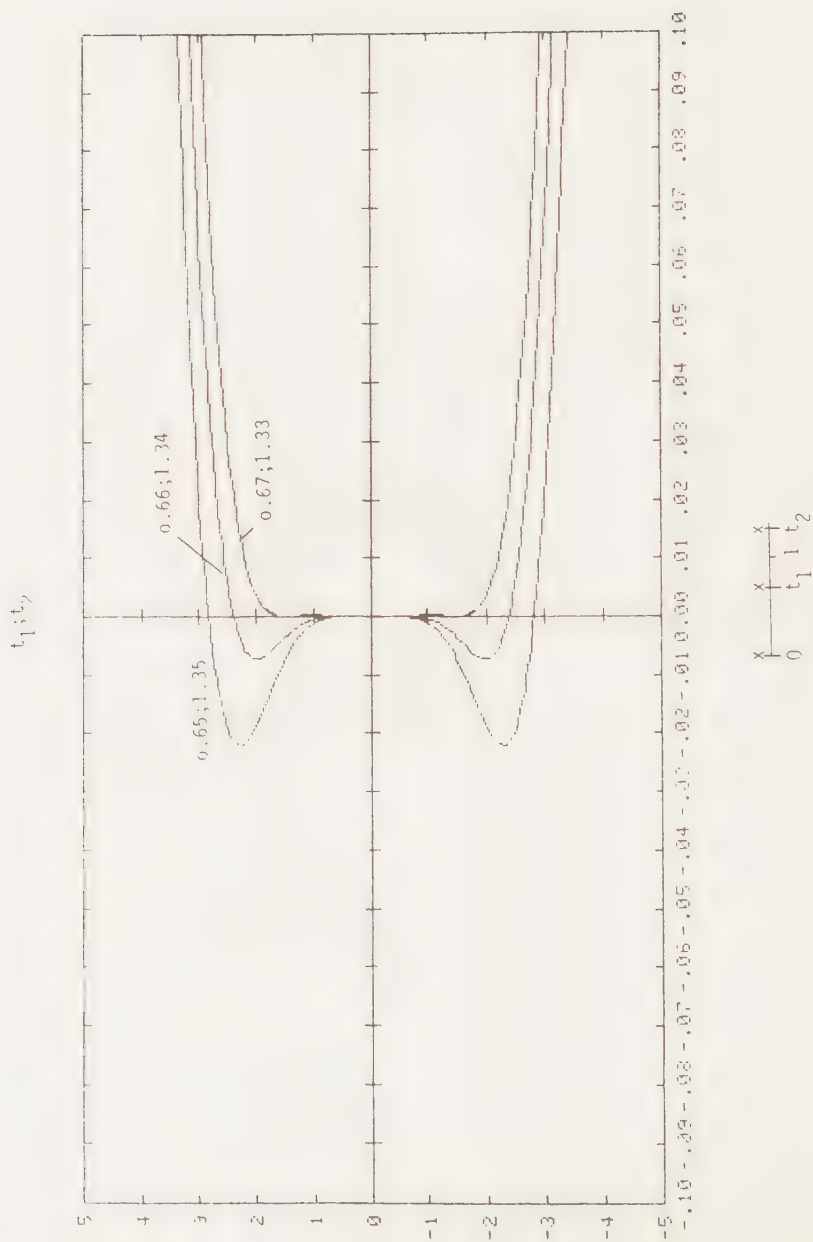
Plot 7



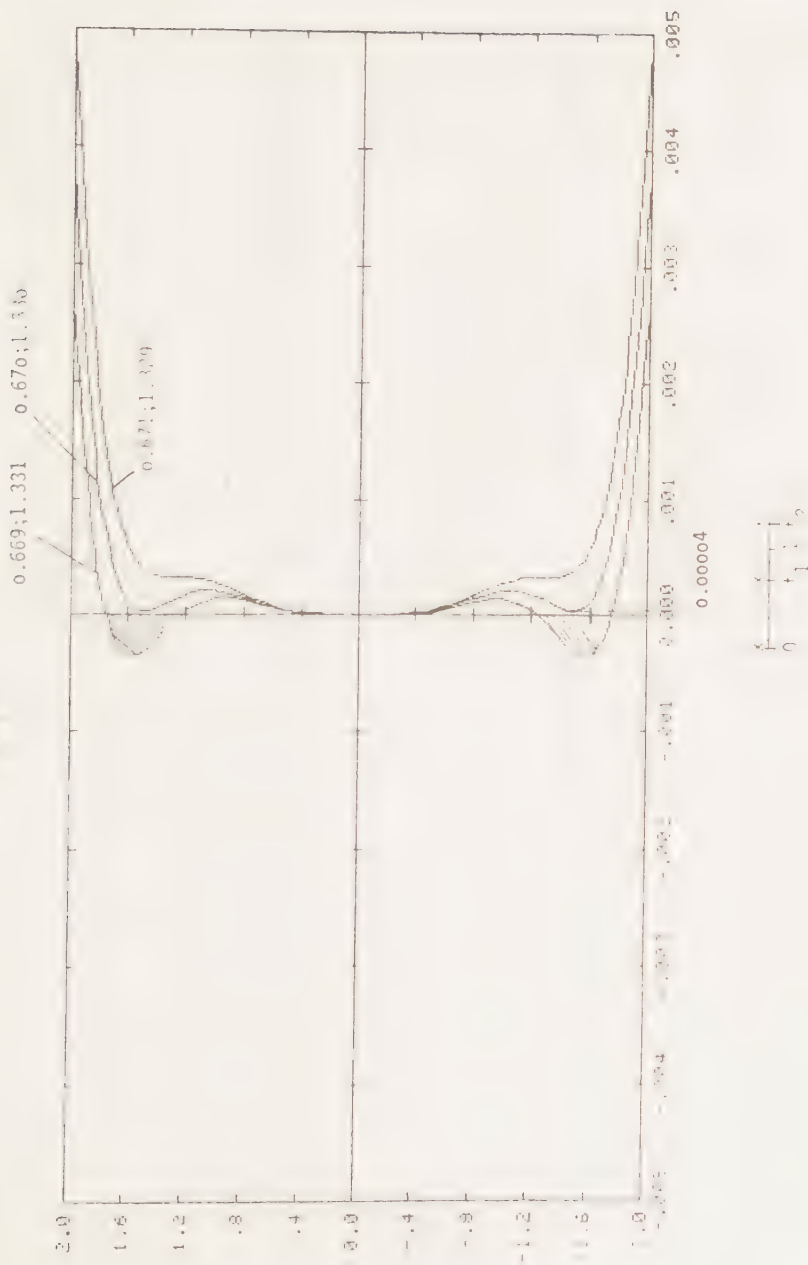
Konfluenz : $\text{Übergang } \partial R_A(\delta) \xrightarrow{\delta \rightarrow 0} \partial R_A(0.0000) = \partial \tilde{R}_A$



Plot 9

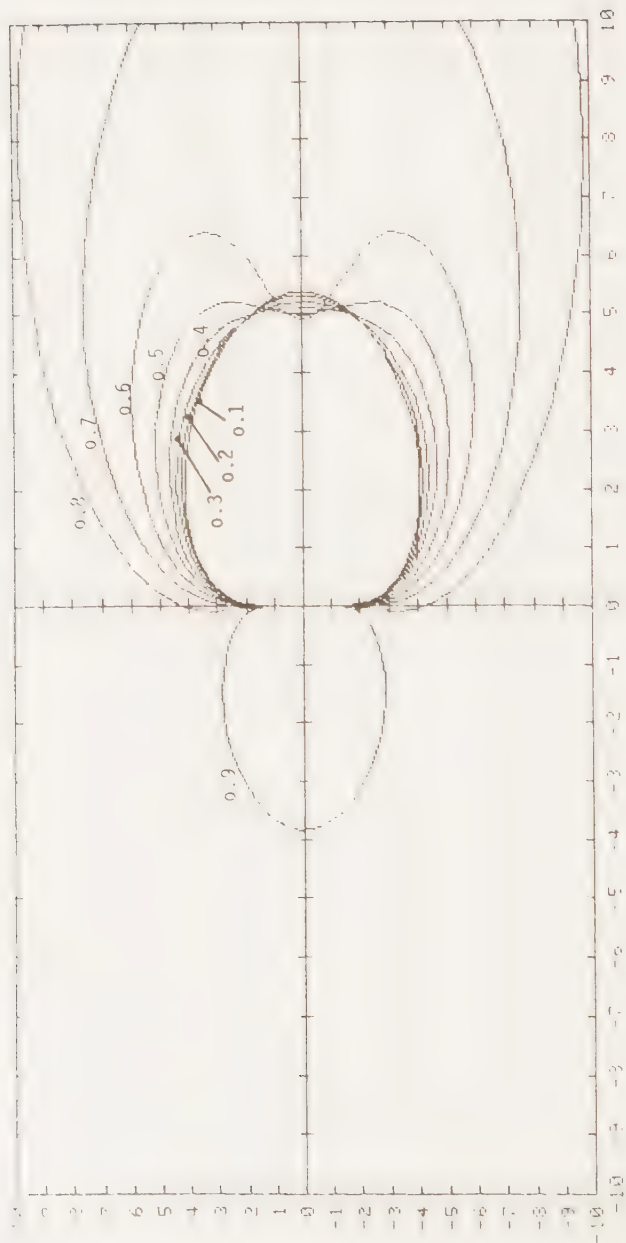


Plot 10



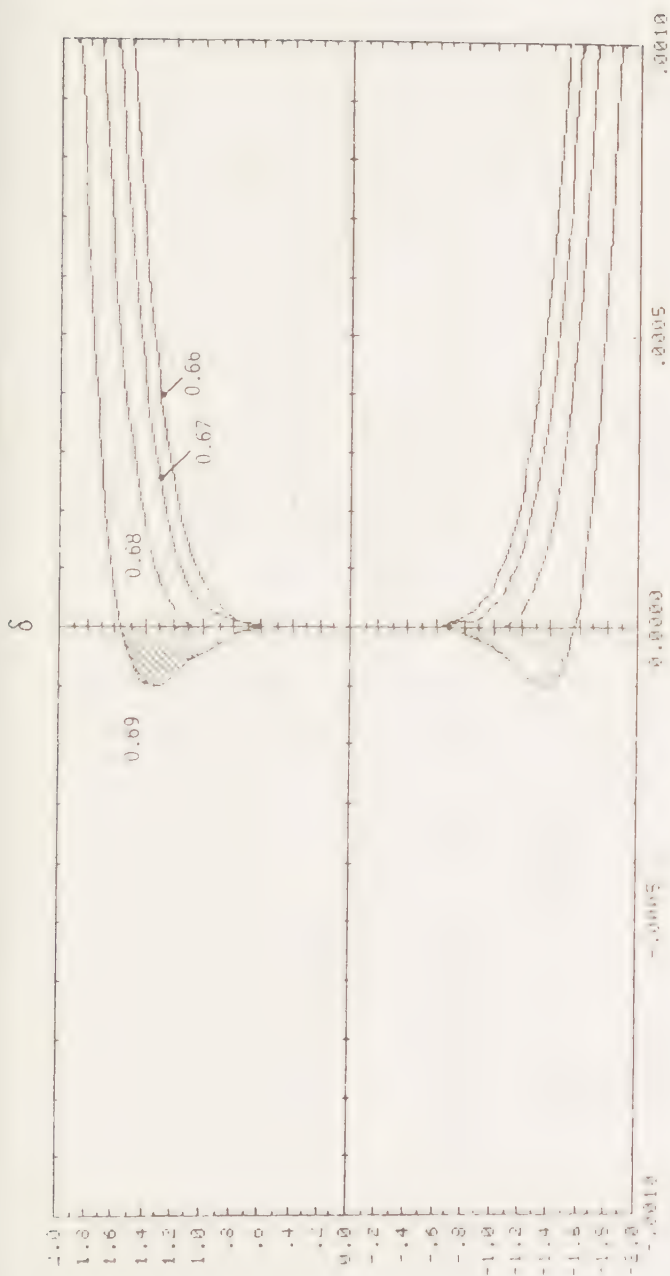
Plot 11

δ



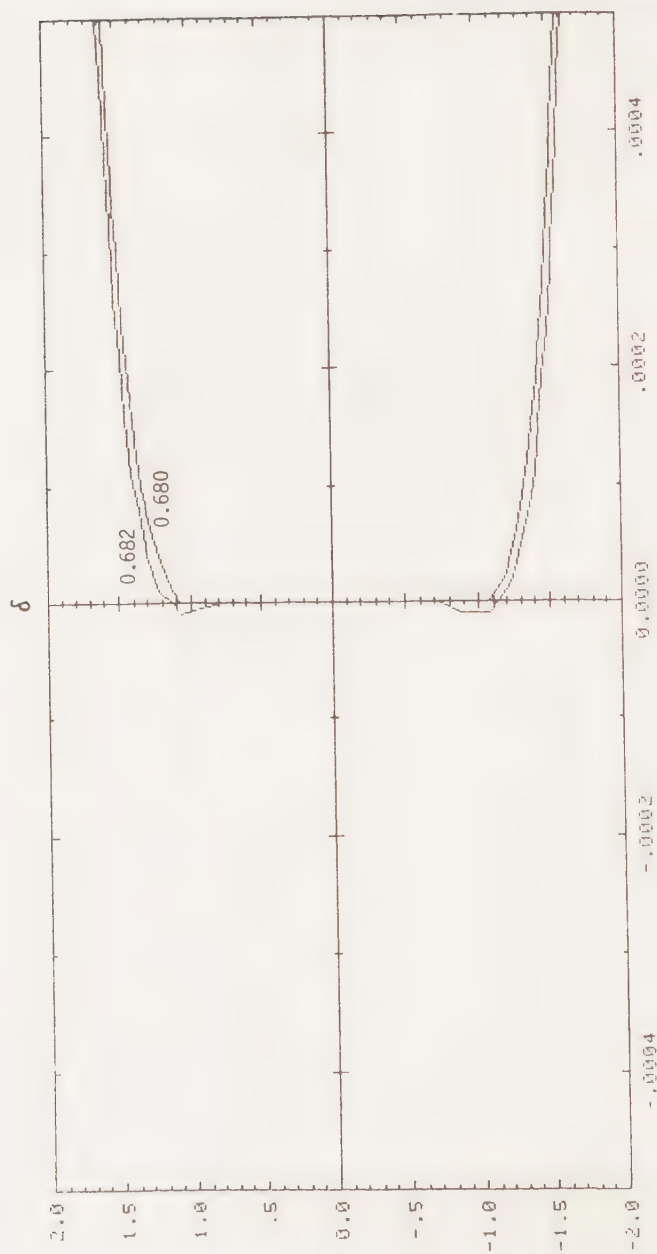
$$\begin{matrix} \delta \\ 0 \end{matrix} \begin{matrix} \delta \\ t_1 \end{matrix} \begin{matrix} \delta \\ 1 \end{matrix} \begin{matrix} \delta \\ t_2 \end{matrix} \begin{matrix} \delta \\ t_3 \end{matrix}$$

$$\begin{aligned} t_1 &:= 1 - \delta \\ t_2 &:= 1 + 0.5(1 - \frac{1}{\sqrt{2}})\delta \\ t_3 &:= 1 + 0.5(1 + \frac{1}{\sqrt{2}})\delta \end{aligned}$$

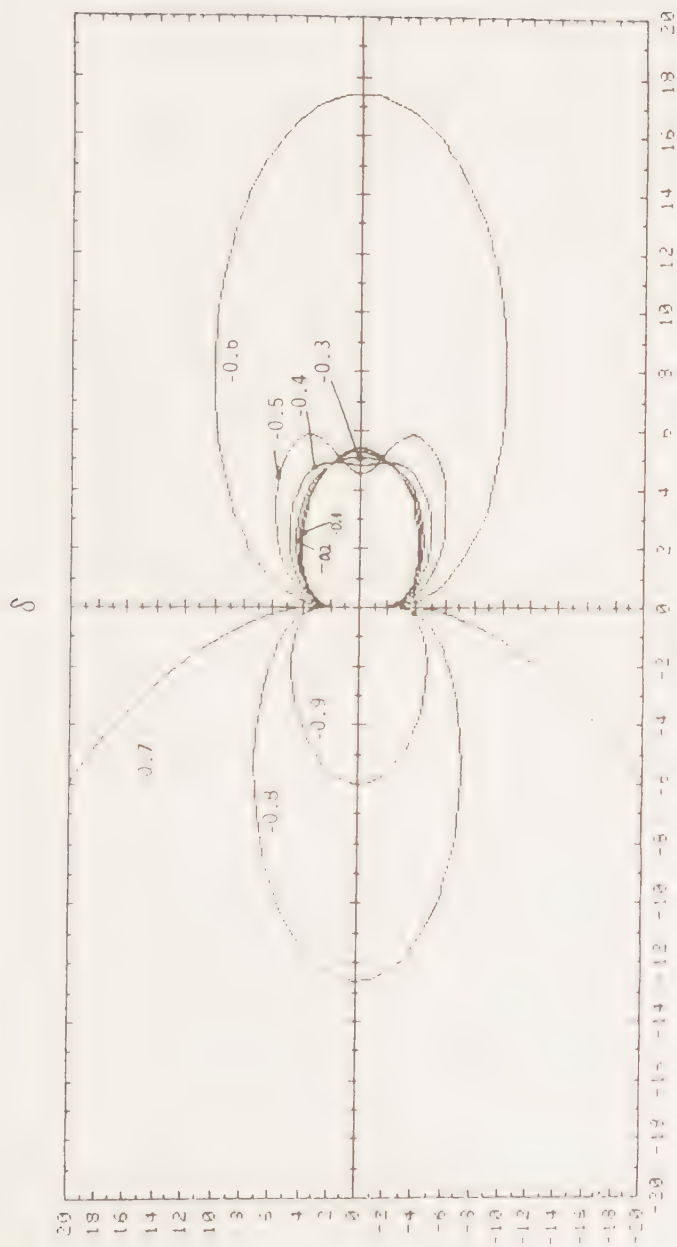


gleiche Parameterkonfiguration wie bei Plot 11

Plot 13

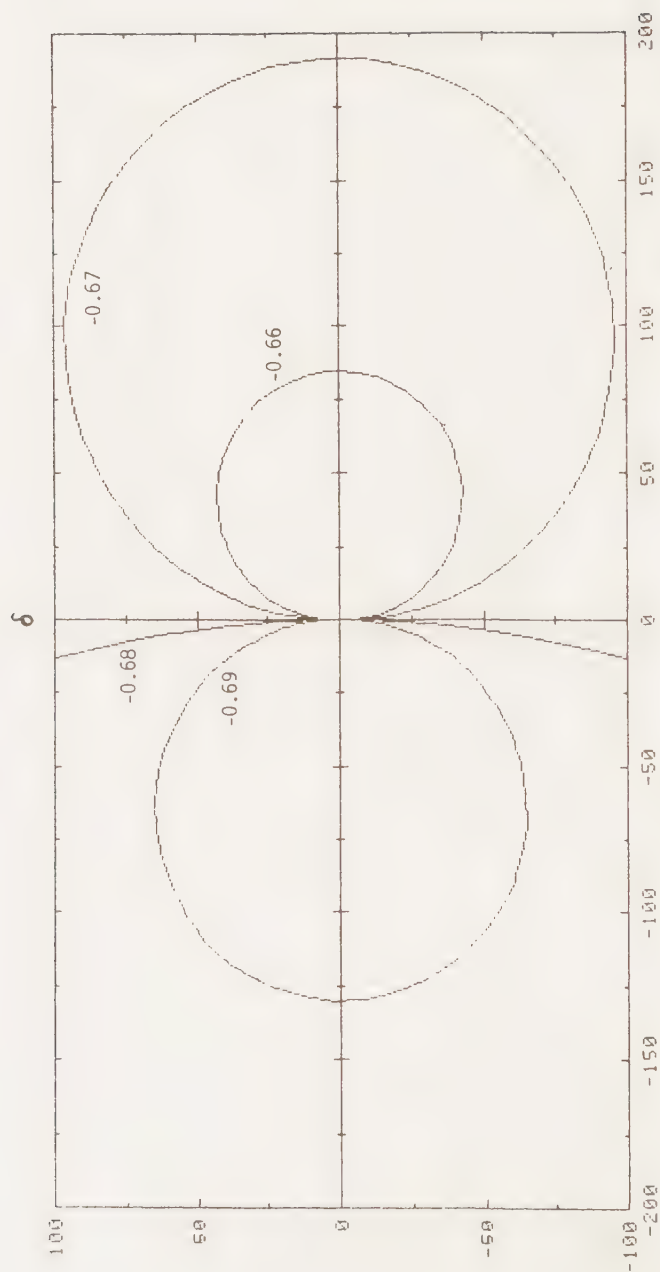


gleiche Parameterkonfiguration wie bei Plot 11



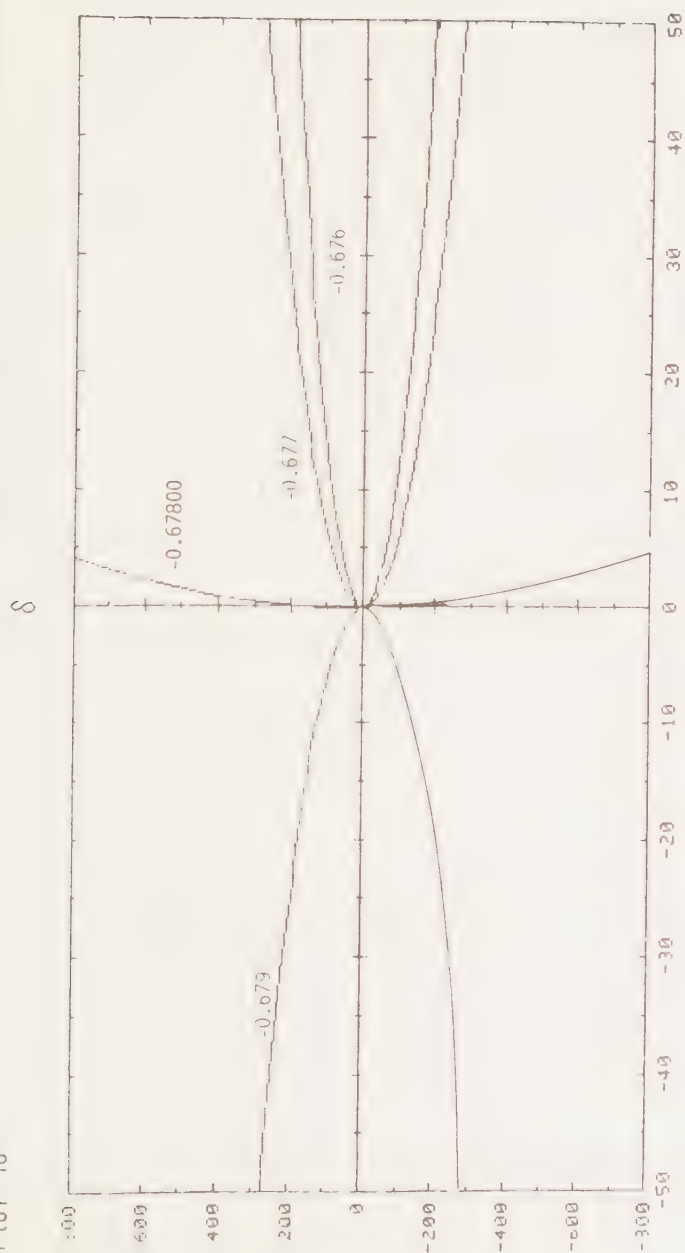
gleiche Parameterkonfiguration wie bei Plot 11

Plot 15



gleiche Parameterkonfiguration wie bei Plot 11

Plot 16



gleiche Parameterkonfiguration wie bei Plot 11

Literaturverzeichnis

- [1] Burrage, K., A special family of Runge-Kutta methods for solving stiff differential equations, BIT 18(1978), 22-41
- [2] Burrage, K./ Chipman, F.H., The stability properties of singly-implicit general linear methods, Auckland Computer Science Reports No.30, June, 1983
- [3] Butcher, J.C., A generalization of singly-implicit methods, BIT 21(1981), 175-189
- [4] Dahlquist, G.G., Convergence and stability in the numerical integration of ordinary differential equations, Math. Scand.4 (1956), 33-53
- [5] Dahlquist, G.G., A special stability problem for linear multistep methods, BIT 3(1963), 27-43
- [6] Friedli, A./ Jeltsch, R., An algebraic test for A_0 -stability, BIT 18(1978), 402-414
- [7] Fuchs, P.M., Numerische Behandlung steifer Anfangswertprobleme, Diplomarbeit Mainz, 1979
- [8] Fuchs, P.M., Charakterisierung der Konvergenzordnung und A_∞ -Stabilität von speziellen Kollokationsverfahren, ZAMM 63(1983), T345-T346
- [9] Gear, C.W., The automatic integration of stiff ordinary differential equations, Proceedings of IFIP Congress 1968, North Holland Publishing Company, Amsterdam (1969), 187-193
- [10] Genin, Y., An algebraic approach to A-stable multistep multiderivative integration formulas, BIT 14(1974), 382-406
- [11] Grigorieff, R.D., Numerik gewöhnlicher Differentialgleichungen 2, Teubner Studienbücher Mathematik, Stuttgart 1977
- [12] Hairer, E., Constructive characterization of A-stable approximations to $\exp(z)$ and its connection with algebraically stable Runge-Kutta methods, Numer. Math. 39(1982), 247-258
- [13] Householder, A.S., The theory of matrices in numerical analysis, Blaisdell Publishing Company, New York.Toronto.London 1964
- [14] Hurwitz, A., Über die Nullstellen der Besselschen Funktion, Math. Ann. 33(1889), 246-266

- [15] Jeltsch, R., Note on A-stability of multistep multiderivative methods, BIT 16(1976), 74-78
- [16] Jeltsch, R., Stiff stability of multistep multiderivative methods, SIAM J.Numer.Anal. 14(1977), 760-772
- [17] Jeltsch, R., Stability on the imaginary axis and A-stability of linear multistep methods, BIT 18(1978), 170-174
- [18] Jeltsch, R., Corrigendum: Stiff stability of multistep multiderivative methods, SIAM J.Numer.Anal. 16(1979), 339-345
- [19] Jeltsch, R., A_0 -stability and stiff stability of Brown's multistep multiderivative methods, Numer.Math. 32(1979), 167-181
- [20] Jeltsch, R./ Nevanlinna, O., Stability of explicit time discretizations for solving initial value problems, Numer. Math. 37(1981), 61-91
- [21] Jeltsch, R./ Nevanlinna, O., Stability and accuracy of time discretizations for initial value problems, Numer.Math. 40(1982), 245-296
- [22] Kühle, U., Stabilitätsgebiete der extrapolierten Sehnentrapezregel bei gewöhnlichen Differentialgleichungen, Diplomarbeit Köln, 1976
- [23] Mülthei, H.N., Splineapproximationen von beliebigem Defekt zur numerischen Lösung gewöhnlicher Differentialgleichungen. Teil I, Numer.Math. 32(1979), 147-157
- [24] Mülthei, H.N., dto.Teil II, Numer.Math.32(1979), 343-358
- [25] Mülthei, H.N., dto.Teil III, Numer.Math.34(1980), 143-154
- [26] Mülthei, H.N., Numerische Lösung gewöhnlicher Differentialgleichungen mit Splinefunktionen, Computing 25(1980), 317-335
- [27] Mülthei, H.N., Zur numerischen Lösung gewöhnlicher Differentialgleichungen mit Splines in einem Sonderfall, Math. Meth. in the Appl.Sci. 2(1980), 419-428
- [28] Mülthei, H.N., Maximale Konvergenzordnung bei der numerischen Lösung von Anfangswertproblemen mit Splines, Numer.Math. 39(1982), 449-463
- [29] Mülthei, H.N., A-stabile Kollokationsverfahren mit mehrfachen Knoten, Computing 29(1982), 51-61

- [30] Nørsett, S.P., C-polynomials for rational approximations of the exponential function, Numer.Math. 25(1975), 39-56
- [31] Nørsett, S.P., Restricted Padéapproximations to the exponential function, SIAM J.Numer.Anal. 15(1978), 1008-1029
- [32] Nørsett, S.P./ Wanner, G., The real-pole sandwich for rational approximations and oscillation equations, BIT 19(1979), 79-94
- [33] Nørsett, S.P./ Wolfbrandt, A., Attainable order of rational approximations to the exponential function with only real poles, BIT 17(1977), 200-208
- [34] Obreschkoff, N., Verteilung und Berechnung der Nullstellen reeller Polynome, VEB Deutscher Verlag der Wissenschaften, Berlin 1963
- [35] Ostrowski, A.M., Solution of equations in Euclidian and Banach spaces, Third Edition, Academic Press, New York. London, 1973
- [36] Routh, E.J., A treatise on the stability of a given state of motion, London 1877
- [37] Saff, E.B./ Varga, R.S., On the zeros and poles Padé approximants to e^z , Numer.Math. 25(1975), 1-14
- [38] Saff, E.B./ Varga, R.S., dto. II, Padé and rational approximations: Theory and applications, 195-213, New York: Academic Press 1977
- [39] Saff, E.B./ Varga, R.S., dto. III, Numer.Math. 30(1978), 241-266
- [40] Schneider, C., Generalized singly-implicit Runge-Kutta methods with arbitrary knots, zur Veröffentlichung einge-reicht
- [41] Stoer, J./ Bulirsch, R., Einführung in die numerische Mathematik II, Heidelberger Taschenbücher 114, Springer-Verlag, Berlin.Heidelberg.New York 1973
- [42] Wanner, G., On the choice of γ for singly-implicit RK or Rosenbrock methods, BIT 20(1980), 102-106
- [43] Wanner, G., Characterization of all A-stable methods of order $2m-4$, BIT 20(1980), 367-374

- [44] Wanner, G., On Nordsieck-type collocation methods or spline-collocation methods or generalized singly-implicit RK methods, unveröffentlicht
- [45] Wanner, G./ Hairer, E./ Nørsett, S.P., Order stars and stability theorems, BIT 18(1978), 475-489
- [46] Werner, W., Beiträge zur Numerik nichtlinearer Gleichungen, Habilitationsschrift Mainz, 1983
- [47] Gautschi, W., On the inverses of Vandermonde and confluent Vandermonde matrices, Numer.Math. 4(1962), 117-123

Lebenslauf

Personalien

Name: Fuchs, Peter Michael
Geburt: 23. April 1955 in Frankfurt/Main
Eltern: Dr. Ernst-Günter Fuchs, Dipl.-Chem.
Ursula Fuchs, geb. Vogel

Schulischer Werdegang

1961 - 1965 Karl von Ibell-Schule in F-Unterliederbach
1965 - 1972 Leibniz-Gymnasium in F-Höchst
1973, 18. 1. Reifeprüfung

1974, 2. 1. Beginn der Mitgliedschaft beim Katastrophenschutz des Main-Taunus Kreises, Technisches Hilfswerk, OV Hofheim

Studium

1973, 20. 3. Immatrikulation an der Johannes Gutenberg Universität in Mainz für die Fächer Mathematik und Physik
1975, 20.10. Diplomvorprüfung
1979, 3. 5. Diplomprüfung in Mathematik, Nebenfach Physik

1982, 1.10. Beginn der Tätigkeit als wissenschaftlicher Mitarbeiter am Fachbereich Mathematik der Johannes Gutenberg Universität in Mainz

6233 Kelkheim (Taunus), den 23. Oktober 1984

Peter Fuchs

